THESIS

TROPICAL CYCLONE TRACK PREDICTABILITY AND THE ADJOINT METHOD OF DATA ASSIMILATION

Submitted by

Bruce W. Thompson

Department of Atmospheric Science

In partial fulfillment of the requirements

for the degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 1989

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS NONE
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFIT/CI/CIA- 89-065
6a. NAME OF PERFORMING ORGANIZATION AFIT STUDENT AT COLORADO STATE UNIVERSITY 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFIT/CIA
6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code) Wright-Patterson AFB OH 45433-6583
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification) (UNCLASSIFIED)
Tropical Cyclone Track Predictability and the Adjoint Method of Data Assimilation

12. PERSONAL AUTHOR(S)
Bruce W. Thompson

13a. TYPE OF REPORT THESIS/~~DISSERTATION~~	13b. TIME COVERED FROM _____ TO _____	14. DATE OF REPORT (Year, Month, Day) 1989	15. PAGE COUNT 78

16. SUPPLEMENTARY NOTATION
APPROVED FOR PUBLIC RELEASE IAW AFR 190-1
ERNEST A. HAYGOOD, 1st Lt, USAF
Executive Officer, Civilian Institution Programs

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL ERNEST A. HAYGOOD, 1st Lt, USAF	22b. TELEPHONE (Include Area Code) (513) 255-2259 22c. OFFICE SYMBOL AFIT/CI

DD Form 1473, JUN 86 *Previous editions are obsolete.* SECURITY CLASSIFICATION OF THIS PAGE

AFIT/CI "OVERPRINT"

COLORADO STATE UNIVERSITY

May 11, 1989

WE HEREBY RECOMMEND THAT THE THESIS ENTITLED TROPICAL CYCLONE TRACK PREDICTABILITY AND THE ADJOINT METHOD OF DATA ASSIMILATION PREPARED UNDER OUR SUPERVISION BY BRUCE W. THOMPSON BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE.

Committee on Graduate Work

Committee Member

Committee Member
Gerald D. Taylor

Adviser
Wayne H. Schubert

Department Head
Stephen K. Cox

ABSTRACT OF THESIS

TROPICAL CYCLONE TRACK PREDICTABILITY AND THE ADJOINT METHOD OF DATA ASSIMILATION

In this study, we explore a new method to improve hurricane track forecasts. This is done by modifying the model's initial conditions using the adjoint method developed by Talagrand and Courtier (1987). The idea is to run the model forward using the governing equation, and then run the model backwards using the adjoint equation. The result of the forward integration is the distance function, and the result of the backward integration is the gradient of the distance function, where the distance function is a scalar measure of the distance between the observed and model hurricane track. The gradient of the distance function is used in a minimization scheme that modifies the initial conditions. These new initial conditions produce a model track closer to the observed track.

Like Talagrand and Courtier, we derive the adjoint method using the spectral non-divergent vorticity equation. However, to eliminate computational error, here we use the Adams-Bashforth time integration scheme instead of the leapfrog method. Experiments were run using the nondivergent barotropic model to indicate how the adjoint method can improve hurricane track forecasts. First, the model was integrated forward using slightly different vortices and their results compared. The results show that small changes in the vortex produce large changes in the vortex track. This indicates that there is important information in the vortex track itself, and that the adjoint method can be very useful in improving the track forecast. Then, we ran experiments that showed that the vortex

track is very sensitive to changes in the vortex structure. These experiments show that subtle modification of the vortex using the adjoint method can improve hurricane track predictions.

Bruce W. Thompson
Department of Atmospheric Science
Colorado State University
Fort Collins, Colorado 80523
Summer 1989

ACKNOWLEDGEMENTS

I would like to thank Professor Wayne H. Schubert for all his support and guidance throughout this work, and also Mr. Paul Ciesielski and Professor Gerald Taylor for their valuable help. I would also like to thank Ms. Jenny Martin for her help in editing and the U.S. Air Force for giving me this opportunity. And a special thanks to my wife Shue-Jane Lee for all her love and support. Funding for this research was provided by the National Science Foundation under Grant ATM 8510664 and its continuation ATM 8814541. Additional funding was provided by ONR Grant N00014-87-K-0535.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

Tropical cyclones are the most significant, and intensely studied phenomena in the tropics. The strong wind, intense convection, and heavy rains associated with these storms cover hundreds of square miles. Naturally, because of the tremendous destructive forces associated with tropical storms, meteorologists concentrated on forecasting their movement. A background of the predictive methods to forecast hurricane movement (Simpson and Riehl, 1981) is discussed below.

1.1 Review of hurricane prediction models

Before the development of dynamic prediction models, hurricane movement was regarded as the response of a vortex to a steering current. So forecasts were made by identifying the steering current and determining how this current and future changes in this current would modify the future track of the system. This process is still pursued by forecasters to test the credibility of model results.

In the 1940's, Grady Norton used the wind direction and speed at the top of the hurricane as an index to the movement of the vortex. However, it was not known to what degree the circulation within the vortex influenced the displacement of the center, so this did not provide sufficiently accurate forecasts.

One of the earliest models to provide objective predictions of movement was proposed by H. Riehl in 1956. Riehl considered the best available index to steering the hurricane was the geostrophic flow of the environment at the level of nondivergence (4-6 km). He computed zonal and meridional components of geostrophic wind from 500 mb analyses using a rectangular grid superimposed on the vortex. This data was used as input to a regression based upon historic storm cases to obtain the westward and northward components of

displacement for the ensuing 24 hour period. The method worked quite well in a research environment, but operationally it suffered from subjectivity of hurried hand-analyses for the 500 mb surface.

During the late 1950's and early 1960's, the search for methods less sensitive to subjective analysis led to the use of statistical screening procedures to select predictors from the surface charts. This produced a model that, although it took no cognizance of the upper-level circulation, displayed good skill with westward-moving hurricanes and blazed the trail for the development of a hierarchy of similar but more powerful models for use at the National Hurricane Center. These incorporated the best conceptual features of the Riehl model and the early screening models. The application of these models, none of which included dynamic and energy processes, was primarily responsible for a significant increase in forecast skills at the National Hurricane Center in the 1960's.

During this same time period the first attempts at predicting tropical cyclone tracks using dynamical models were made (e.g., Sasaki, 1955; Kasahara, 1957). Because of the limited computational resources it was necessary to use simple barotropic models and to treat the small-scale vortex circulation separately from the larger-scale flow. In the 1960's the computing power increased and it was no longer necessary to treat the vortex separately (e.g., Birchfield, 1960; Sanders and Burpee, 1968).

The first completely objective procedure for predicting hurricane movement, using machine analyses of current weather data, was developed for use at the National Hurricane Center in 1964 and is known as NHC-64. This model employed predictors obtained from analyses of circulations at 1000, 700, and 500 mb over a large synoptic-scale domain. The predictors are based upon statistical screening of data from historical hurricane cases.

1.2 Current prediction models

Currently, there are three classes of models used for predicting hurricane movement: 1) statistical analog models, 2) dynamic analog models, and 3) pure dynamical models. The first draws upon the climatology of hurricane tracks and of persistence of movement to produce a most-probable displacement of the center. The output is a function of initial

position, past movement, and calendar time of occurrence. But, the computation does not take into account the environment or its influences on the hurricane. The second class extracts from historical cases the dynamical properties of the near or the large-scale environment that correlate with some aspect of hurricane movement. These are combined in a multiple regression statement as analogs to the migration of the vortex. The third class, not concerned with history, combines basic principles of fluid motion, the thermodynamics of an ideal gas, and the application of conservation relationships to predict the behavior and movement of the vortex.

The first two classes suffer from incomplete hurricane climatology, especially with regard to the cases with critical changes in movement and strength. Dynamical models encounter at least three kinds of problems. The first, and probably the most important, is that of initialization. Due to the lack of data in the tropics the description of the initial state of the atmosphere when computations begin, especially the description of the processes in the vigorous inner core of the vortex, may not be very good. Second, while the environment can be described adequately with a grid spacing at 300 km, the active vortex, especially approaching the region of maximum winds, may require a grid spacing of 5-10 km. Thus, higher horizontal resolution is needed to describe what is going on in the inner core of the hurricane. Currently, a horizontal resolution of 5 to 10 km is too costly in time and resources to be used. Finally, an adequate parameterization of the heating generated by cumulus convection has not been obtained. The current models used by the National Hurricane Center are described below, (Neumann and Pelissier, 1981).

1.2.1 HURRAN

HURRAN (HURRicane ANalog) is a statistical analog model which uses the current position, direction, and speed of movement of the system for the preceding 12 hours to compute a track. Drawing upon past hurricane tracks, the model computes the most probable track for a 72 hour period based on the movement of historic hurricanes that occurred at the same time of the year and whose positions and motion vectors were similar to the present case. The principle shortcoming of HURRAN and other analog methods is

limited usefulness during highly anomalous movements, since too few analogs are available for computing tracks on such occasions.

1.2.2 CLIPER

The CLIPER (CLImatology and PERsistence) model was intended as a backup for HURRAN when the latter failed to provide a forecast because of insufficient analogs. However, CLIPER consistently outperforms HURRAN, particularly on recurving storms. This method draws its predictors solely from climatology and persistence (of past motion). So the CLIPER has the advantage of always providing a forecast, even under anomalous situations.

1.2.3 NHC67 and NHC72

These two models are similar in concept, and updated versions of the NHC64 model previously discussed. The basic difference between these models and the CLIPER model is the additional use of current and 24 hour old 1000, 700, and 500 mb geopotential height data to modify a preliminary forecast based only on climatology and persistence. Although NHC67 and NHC72 use similar methodology, there are important structural differences that lead to different performance characteristics in an operational environment, one of which is more explicit use of climatology and persistence in the form of direct input from the CLIPER model by NHC72.

1.2.4 NHC73

In the early 1970's, a large number of Atlantic storms with anomalous motion characteristics highlighted the inherent inability of the NHC67 and NHC72 models to forecast such motion with acceptable accuracy. This stimulated the development of the statistical-dynamical NHC73 model that incorporates more recent numerical prognoses. Predictors entering the NHC73 model included 1) the output from the CLIPER model; 2) current 1000, 700, and 500 mb gridded analyses; and 3) 24, 36, and 48 hour geopotential height prognoses from the NMC primitive equation model.

1.2.5 SANBAR

The first operationally successful dynamical model was developed at MIT by Fred Sanders and Robert Burpee. This barotropic model, known as SANBAR, computes pressure-weighted mean winds for the layer 1000 to 100 mb, from which stream functions are generated and used as inputs to the predictive model. In the initialization process, the vortex is replaced with an ideal vortex modified to provide initial steering that is consistent with the observed motion of the system.

1.2.6 MFM

The baroclinic Moveable Fine Mesh (MFM) model was first operational in 1976. The physics of this model is generally the same as other primitive equation (PE) models now in operation. However, one of its unique characteristics is the ability of the grid to follow the storms as they move during a forecast. The MFM model also has much finer resolution in both the horizontal and vertical, then other PE models. The horizontal resolution is 60 km, and there are 10 layers in the vertical. Because of the computer time needed, it is necessary to make the total areal coverage much smaller than other existing operational models. Also, because of the observational and computational limitations, an axisymmetric vortex that is qualitatively similar to the hurricane is used.

Even though many numerical models have been developed over the years to forecast tropical cyclone motion, the official forecasts remain somewhat subjective. Because of deficiencies in the initialization, parameterization, and the knowledge of the interactions between the storm and its environment, all of the National Hurricane Center models are used just to provide guidance in forecasting the storm track. A study done by Neumann and Pelissier (1980) concluded that none of the models can be singled out as clearly superior or inferior, with significant mean forecast errors ranging from around 50 n mi at 12 hours to 350–400 n mi at 72 hours. This study also indicated that certain models are better in certain areas and situations. For example, although the NHC73 model performs well for storms initially located poleward of 24.5° north, the model performance south of this latitude is rather poor. The HURRAN model performance is the opposite, with

better forecasts south of 24.5° north, while only NHC73's 72 hour forecast is better than that of climatology and persistence. Although each model has different characteristics, the key point is that none of the models provide forecasts of the desired accuracy. One way that may improve tropical cyclone motion forecasts by dynamical models, is the use of the following method.

1.3 The adjoint method

Suppose we are trying to make a tropical cyclone motion forecast using a dynamical model. The dynamical model could be of any level of complexity—from the nondivergent barotropic model, through the divergent barotropic model, to the fully three dimensional primitive equations with parameterized moist physics. In this case the nondivergent barotropic model is used. No matter how simple or complex the model, we are faced with the problem of the model initialization based on sparse observations. Observations in the tropical cyclone are often totally lacking, and thus the vortex is essentially unresolved. We are fortunate if there are even enough observations to adequately define the larger scale "steering flow." Under such circumstances, initialization often involves insertion of a "bogus vortex." The dynamical model sometimes moves this vortex in a direction and at a speed quite different from the observed vortex. The track forecast might be improved by changing certain structure parameters of the bogus vortex, e.g. its size, strength, or tangential asymmetry. However, such procedures remain somewhat arbitrary.

The procedure described in the preceding paragraph makes no use of mass, wind, or track data prior to the initialization time. Now, let us look at the problem using the concept of four dimensional data assimilation, and, in particular, the adjoint method, which results in specifying complete initial conditions at a given instant from observations distributed in space and time. Suppose we have acquired (over the time interval $t_0 \leq t \leq t_P$) large-scale wind and temperature data from scattered island radiosonde stations and a sequence of tropical cyclone position fixes from satellite data. If we are forecasting with the nondivergent barotropic model we only need to consider the vorticity field. Then assume that the cyclone tracks can be somehow converted to "vorticity observations" so

that we have $\zeta_{obs}(x,y,t)$ for $t_0 \leq t \leq t_P$, where x is the east direction and y is the north direction on a beta plane. A model forecast from time t_0 with initial conditions $\zeta_0(x,y)$ produces the field $\zeta(x,y,t)$. Let J be the integral (spatial and temporal) measure of the squared difference between $\zeta(x,y,t)$ and $\zeta_{obs}(x,y,t)$, which is expressed as

$$J = \frac{1}{2}\int_{t_0}^{t_1}\int_0^L\int_0^L \Big(\zeta(x,y,t) - \zeta_{obs}(x,y,t)\Big)^2 dxdydt \tag{1.1}$$

Since $\zeta_{obs}(x,y,t)$ is known and $\zeta(x,y,t)$ is determined from $\zeta_0(x,y)$ by model integration, the "distance function" J depends only on $\zeta_0(x,y)$. How can we vary the $\zeta_0(x,y)$ field to make J as small as possible? Stated differently, how can we massage the data at t_0 in order to make the model track fit closely with the observed track over the interval $t_0 \leq t \leq t_P$? If we could do this we would have a model field close to the observed field over the interval $t_0 \leq t \leq t_P$, and intuition would suggest that continuation of this model solution past t_P would give a pretty good track. At the very least, the model vortex should be going the right direction and speed at t_P.

Let us try to minimize J in a naive and brute force fashion. Consider a discretized model with N^2 degrees of freedom (N by N points, say, where $N = 256$). Thus $\zeta_0(x,y)$ is represented by a vector ζ_0 of length N^2. Let $\nabla_{\zeta_0} J$ denote the gradient of J with respect to each element of ζ_0, which means $\nabla_{\zeta_0} J$ is a vector of length N^2. The first element of $\nabla_{\zeta_0} J$ tells us how the distance function J would change if we modified the initial condition at the first point of the grid, and so on through all the points of the grid. Thus, if we know $\nabla_{\zeta_0} J$, we can make simultaneous, subtle modifications of ζ_0 at all points in order to reduce J. We conclude that the knowledge of $\nabla_{\zeta_0} J$ might give us considerable power to improve track forecasts. The brute force method of determining $\nabla_{\zeta_0} J$ consists of making a forecast using ζ_0 as the initial condition, followed by N^2 more forecasts with ζ_0 slightly modified in turn at each grid point; each of the N^2 forecasts is compared to the original forecast and the associated change in J calculated. Unfortunately, the apparent necessity of making thousands of model runs would probably render the forecast untimely (what's additionally troublesome is that the above procedure has to be iterated).

Here comes the adjoint method to the rescue. The adjoint method can give us $\nabla_{\zeta_0} J$ in a time equivalent to only a few model runs. This is a powerful result, and here is all

we need to do. Derive the tangent linear equation, which in this case is the nondivergent barotropic vorticity equation linearized about the present model solution $\zeta(x, y, t)$. Then find the adjoint of the tangent linear equation. Finally, run the original nonlinear model forward in time from t_0 to t_P, followed by a backward integration from t_P to t_0 using the adjoint equation. If this is done in the proper fashion, the output is $\nabla_{\zeta_0} J$, which can be used to modify ζ_0 and give a better track simulation. The ζ_0 field can be iteratively modified in this same fashion. Since the adjoint model takes about the same computer time as the forecast model, each iteration is roughly equivalent to two forecast runs, and a typical five iteration procedure is equivalent to ten forecast runs.

Although the adjoint method should improve the track forecasts of dynamical models, there are still errors associated with this method. First of all, there are errors with the models themselves, since they can not accurately represent the real atmosphere. In this study we are using a very simple model, the nondivergent barotropic model, which is not very realistic. The errors associated with this model and errors associated with temporal and spatial differencing will discussed in chapter 2 and 3. However, the most significant error, one which was mentioned earlier, is that of initialization. One thing the adjoint method will not improve is errors in the location of the storm center. With the use of satellite data there are errors in the position of the storm center, thus, the speed and direction of the observed storm track could be in error. This is an error that all models encounter and can not be eliminated. There can also be errors associated with the adjoint method which will be discussed later in chapter 2.

The description of the nondivergent barotropic model is in Chapter 2, which includes a discussion on how the adjoint method is integrated into the basic model. The adjoint method and the descent algorithm used in this method are also described in Chapter 2. In Chapter 3, we derive the solution to the nondivergent vorticity equation and the adjoint equation using the Galerkin method and the Adams-Bashforth time differencing scheme. The results from the nondivergent barotropic model using the adjoint method are presented in Chapter 4. The model is run with an axisymmetric initial vortex on a β-plane.

Chapter 2

NONDIVERGENT BAROTROPIC MODEL

2.1 Model description

The governing equation for this model as described in DeMaria (1985) is the conservation of absolute vorticity on a midlatitude β–plane which can be written as

$$\frac{\partial \zeta}{\partial t} + \frac{\partial}{\partial x}(u\zeta) + \frac{\partial}{\partial y}(v\zeta) + \beta v = 0, \tag{2.1}$$

where ζ is the vertical component of the relative vorticity, u and v the eastward and northward components of the nondivergent wind, and β the northward gradient of the Coriolis parameter. In (2.1), ζ can be related to the horizontal wind components by introducing a streamfunction ψ where

$$\zeta = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2}, \tag{2.2}$$

$$u = -\frac{\partial \psi}{\partial y}, v = \frac{\partial \psi}{\partial x}. \tag{2.3}$$

Equations (2.1)–(2.3) were solved using a spectral method with Fourier basis functions on a doubly-periodic domain. For this purpose, the dependent variables are expanded in double Fourier series, and substituted into (2.1)– (2.3) and the adjoint of the linearized version of (2.1). The standard Galerkin procedure is then applied which gives equations for the time dependent series amplitudes. The spectral form of (2.1) and its adjoint are solved using the second order Adams-Bashforth time differencing scheme. A more detailed description of these numerical methods are in sections 3.1 and 3.2. The model was truncated at wavenumber 47 in the x and y directions on a 4000 km square domain. The shortest wave in the model then has a wavelength of about 85 km.

The model uses an axisymmetric modified exponential vortex defined by

$$V = V_m \left(\frac{2(r/r_m)}{1 + (r/r_m)^2} \right) \exp[-a(r/r_m)^b], \tag{2.4}$$

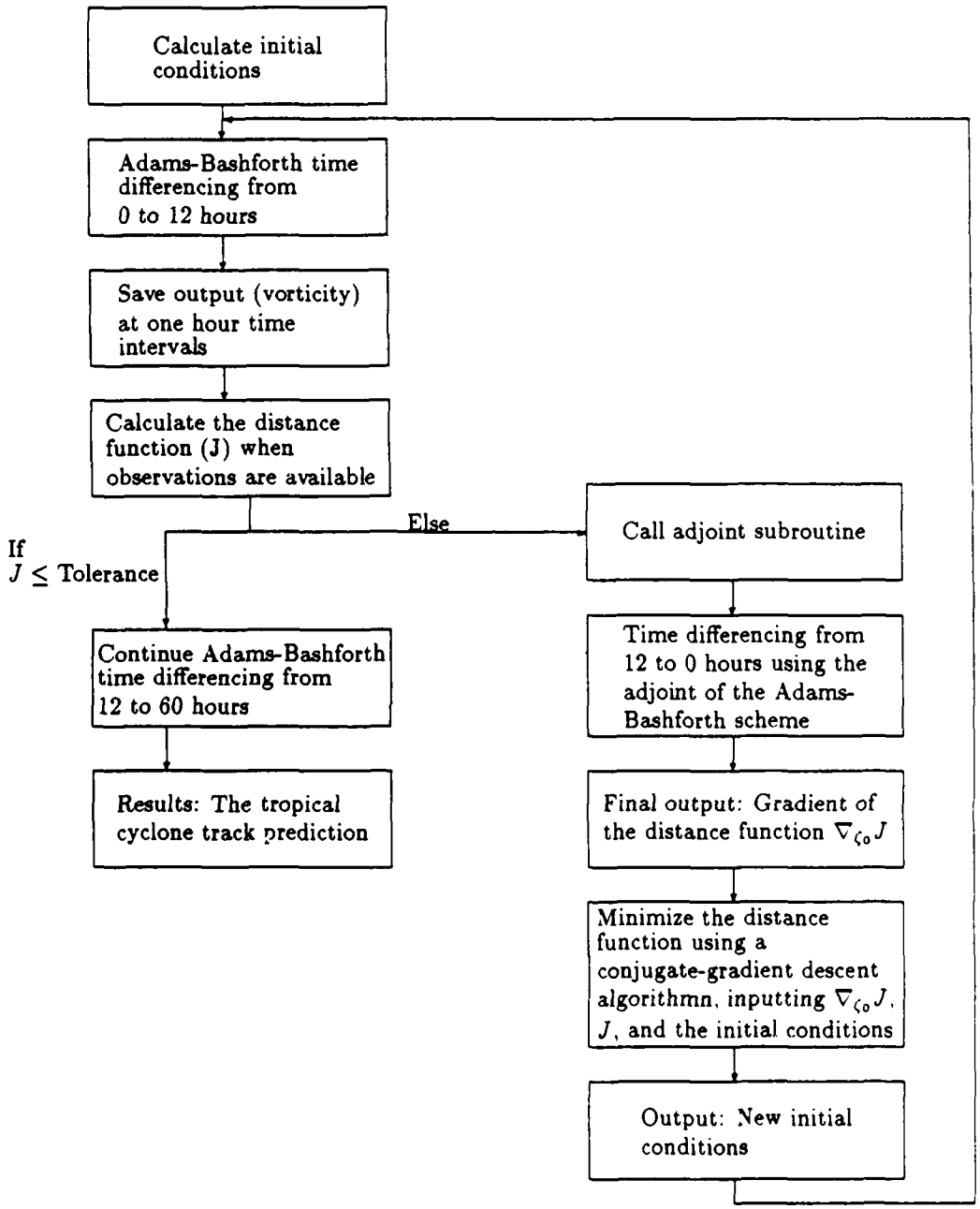
where V is the tangential wind, r the radial distance from the vortex center, V_m the approximate maximum tangential wind (exact for $a = 0$), and r_m the approximate radius of the maximum wind. The exponential factor in equation (2.4) is added so that V decays rapidly with r at large radii. The initial values of V_m and r_m are 30 ms^{-1} and 80 km, respectively, with $a = 10^{-6}$, and $b = 6$. DeMaria used these parameters because they result in a tangential wind profile which lies between the observed profiles for large and small hurricanes. The large scale flow superimposed on this vortex is a zonal flow with a maximum speed of 10 ms^{-1} which represents the tropical, and sub-tropical wind field.

The adjoint method can be inserted into the nondivergent barotropic model in the following way. The original model is first run forward from 0–12 hours, with the output saved at times when observations are available. At 12 hours the adjoint subroutine is called and the adjoint of the linearized vorticity equation is run backwards from 12–0 hours, using the output saved in the forward run. The result of this procedure is the gradient of the distance function, which is then used in a descent algorithm to modify the initial values. If the gradient of the distance function is sufficiently small, the initial conditions will no longer be modified, and the model will continue forward from 12 to 60 hours. Otherwise, the new initial conditions will be used to calculate a new value of the gradient of the distance function. A flowchart of this process is shown in Fig 2.1.

2.2 The adjoint method

The purpose of the adjoint method is to find a solution of an assimilating model which minimizes a given scalar function measuring the "distance " between a model solution and the available observations. This is accomplished by using the adjoint method to compute the gradient of the distance function with respect to the initial conditions. The model is integrated forward in time, where observations are available. This is followed

Flowchart of the model using the adjoint method



by a backwards integration using the adjoint equation. The results of successive forward-backward integrations are gradients of the distance function, which are introduced into a descent algorithm in order to determine the initial conditions which define the minimizing model solution.

In order to do this, the local "tangent" linear equations are used to compute the evolution of the forecast error covariances. The tangent linear equations, given an initial perturbation imposed on a model solution, describe the temporal evolution of the perturbation to the second order in perturbation amplitude. The problem with the tangent linear equation is the ablility to describe the short-term evolution of the forecast error which is at the same time simple enough to be usable in operational practice and accurate enough to improve the quality of present assimilation algorithms.

Instead of performing successive analyses in the course of one integration of the assimilating model, each individual observation being used once without feedback, an alternative approach is used. This is to adjust globally, one model solution to the complete set of available observations. The big advantage is that of exact consistency between dynamics of the model and the final results of assimilation. Then repeat the integrations and correct the model so there is convergence towards the solution, which is compatible to a certain accuracy with the observations.

A more systematic and rigorous approach for globally adjusting one nonlinear model solution to a set of observations distributed in time is the following:

1. Define the scalar function which measures the distance between the solution and the observations.

2. Try to determine the particular solution which minimizes that scalar function. This is a constrained variational problem where the unknowns must minimize a given scalar function while verifying a given set of constraints.

If one notes that a solution of the model is uniquely defined by the corresponding initial state at a given time, the variational problem considered above can be stated in the following terms: find the initial state such that the corresponding model solution minimizes

the scalar function measuring the distance to the observations. The numerical dimensions of the problem are greatly reduced since the minimization process is now performed on the initial state only, and no longer on the whole temporal history of the model. To express the relationship between the variations of the initial state and the corresponding variations of the distance function in a usable form, requires the computation of the gradient of the distance function with respect to the parameters which define the initial state. A possible way to determine this gradient is to perturb in turn all components of the initial state and, for each perturbation, to integrate explicitly the model equations and to compute the corresponding variation of the distance function. But the high numerical cost of this approach probably makes this impractical.

A more practical approach to this problem is using techniques of "optimal control" or more specifically the "adjoint equations" of the assimilating model (Lions 1971). The theory of optimal control deals with the general problem of how the output parameters of a complicated numerical process can be controlled by acting on the input parameters of the process. Among the various tools of optimal control, the adjoint of the local tangent linear equations provide an efficient way for numerically computing the local gradient of a complicated compound function of a set of arguments. In the context of variational data assimilation, the adjoint of the local tangent linear equations of the assimilating model can be used for computing the gradient of the distance function with respect to the model's initial conditions. This gradient is then used for performing a "descent step" in the space of initial conditions, and the process is iterated until some satisfactory approximation of the initial conditions, which minimizes the distance function has been obtained.

2.3 General principle of adjoint theory

There are two properties of a Hilbert space that are used in this theory. A more detailed description of Hilbert space is not needed in the derivation of this theory, but can be found in Reed and Simon (1980).

(i) Let $\mathcal{E}$ be a Hilbert space, with inner product denoted by $\langle , \rangle$, and $\mathbf{v} \rightarrow J(\mathbf{v})$, a differential scalar function defined on $\mathcal{E}$. At any point in $\mathcal{E}$ the differential δJ of J

can be expressed as

$$\delta J = \langle \nabla_{\mathbf{v}} J, \mathbf{v} \rangle, \tag{2.5}$$

where $\nabla_{\mathbf{v}} J$ is the gradient of J with respect to $\mathbf{v}$.

(ii) Let L be a continuous linear operator from $\mathcal{E}$ into $\mathcal{E}$. There exists a unique continuous linear operator L^* also mapping $\mathcal{E}$ into $\mathcal{E}$, with the property that

$$\langle \mathbf{v}, L\mathbf{u} \rangle = (L^*\mathbf{v}, \mathbf{u}). \tag{2.6}$$

L^* is called the adjoint operator of L. In numerical models $\mathbf{u}$ and $\mathbf{v}$ are both vectors, and the matrix which represents L^* is the transpose of the matrix which represents L.

The adjoint theory can be used in numerical models in the following way. Consider a differentiable function $\mathbf{u} \rightarrow \mathbf{v} = \mathcal{G}(\mathbf{u})$, where $\mathbf{u}$ will denote the initial conditions from which a numerical model of the atmospheric flow is integrated, and $\mathbf{v}$ denotes the time sequence of successive model states produced by the integration. The scalar function $J(\mathbf{v})$ will be the scalar function which measures the distance between $\mathbf{v}$ and the available observations, assumed to be distributed over a time interval $[t_0, t_P]$. The model evolution equation is written as

$$\frac{d\mathbf{x}}{dt} = F(\mathbf{x}), \tag{2.7}$$

where the state vector $\mathbf{x}$ belongs to a Hilbert space $\mathcal{E}$ whose inner product is denoted by $\langle\ ,\ \rangle$, and F is a regular function on $\mathcal{E}$. The initial condition $\mathbf{x}(t_0) = \mathbf{u}$ defines a unique solution $\mathbf{x}(t)$ to (2.7). The tangent linear equation at $\mathbf{x}(t)$ can be expressed as

$$\frac{d\delta\mathbf{x}}{dt} = F'(t)\delta\mathbf{x}, \tag{2.8}$$

where $F'(t)$ is the linear operator obtained by differentiating F with respect to $\mathbf{x}$, and $\delta\mathbf{x}$ the perturbation of $\mathbf{x}$. The adjoint of equation (2.8) is

$$-\frac{d\delta\mathbf{x}'}{dt} = F'^*(t)\delta\mathbf{x}', \tag{2.9}$$

whose variable $\delta\mathbf{x}'$ is the adjoint of the perturbation of $\mathbf{x}$ and also belongs to $\mathcal{E}$. Also $F'^*(t)$ is the adjoint of $F'(t)$. The distance function J will be taken as

$$J = \sum_{p=0}^{P} H_p(\mathbf{x}_p), \tag{2.10}$$

where $H_p(\mathbf{x}_p)$ is a scalar measuring the distance between the model and the observations available at time t_p. In what follows, we shall express $H_p(\mathbf{x}_p)$ as

$$H_p(\mathbf{x}_p) = \frac{1}{2}\langle \mathbf{x}(t_p) - \hat{\mathbf{x}}(t_p), \mathbf{x}(t_p) - \hat{\mathbf{x}}(t_p)\rangle, \tag{2.11}$$

where $\mathbf{x}(t_p)$ is the model's solution at time t_p, and $\hat{\mathbf{x}}(t_p)$ are observations at time t_p. Now, the first order variation of J can be expressed as

$$\delta J = \sum_{p=0}^{P} \langle \nabla_{\mathbf{x}} H_p, \delta\mathbf{x}_p\rangle, \tag{2.12}$$

where $\delta\mathbf{x}(t)$ is the first-order variation of $\mathbf{x}(t)$ resulting from a perturbation $\delta\mathbf{u}$ of $\mathbf{u}$ and the gradient of $H_p(\mathbf{x}_p)$ is

$$\nabla_{\mathbf{x}} H_p = \mathbf{x}(t_p) - \hat{\mathbf{x}}(t_p). \tag{2.13}$$

The variation of $\delta\mathbf{x}(t)$ is obtained from $\delta\mathbf{u}$ by integrating (2.8) relative to the solution $\mathbf{x}(t)$. Since (2.8) is linear, the solution at a given time depends linearly on the initial conditions, and can be expressed as

$$\delta\mathbf{x}(t) = R(t, t_0)\delta\mathbf{u}, \tag{2.14}$$

where $R(t, t_0)$ is a well defined linear operator, called the resolvent of (2.8) between times t_0 and t. The resolvent $R(t, t')$ is defined more generally for any two times t and t' lying between t_0 and t_P and possesses the following two properties:

$$R(t, t) = I \qquad \text{for any } t, \tag{2.15}$$

where I is the identity operator on $\mathcal{E}$; and

$$\frac{\partial}{\partial t} R(t, t') = F'(t)R(t, t'), \tag{2.16}$$

for any t and t'. Equation (2.12) can now be rewritten as

$$\delta J = \sum_{p=0}^{P} \langle \nabla_{\mathbf{x}} H(t_p), R(t_p, t_0)\delta \mathbf{u} \rangle. \tag{2.17}$$

Using (2.6), this equation can be expressed as

$$\delta J = \sum_{p=0}^{P} \langle R^*(t_p, t_0)\nabla_{\mathbf{x}} H(t_p), \delta \mathbf{u} \rangle, \tag{2.18}$$

where $R^*(t, t_0)$ is the adjoint of $R(t, t_0)$. From this, $\nabla_{\mathbf{u}} J$ can be shown as

$$\nabla_{\mathbf{u}} J = \sum_{p=0}^{P} R^*(t_p, t_0)\nabla_{\mathbf{x}} H(t_p). \tag{2.19}$$

The adjoint equation (2.9) is also linear, and we denote $S(t', t)$ as its resolvent between times t and t'. For any two solutions $\delta \mathbf{x}(t)$ and $\delta \mathbf{x}'(t)$ of the tangent linear and adjoint equations, (2.8) and (2.9) respectively, the inner product $\langle \delta \mathbf{x}(t), \delta \mathbf{x}'(t) \rangle$ is constant with time since

$$\begin{aligned} \frac{d}{dt}\Big\langle \delta \mathbf{x}(t), \delta \mathbf{x}'(t) \Big\rangle &= \Big\langle \frac{d\delta \mathbf{x}}{dt}(t), \delta \mathbf{x}'(t) \Big\rangle + \Big\langle \delta \mathbf{x}(t), \frac{d\delta \mathbf{x}'}{dt}(t) \Big\rangle \\ &= \Big\langle F'(t)\delta \mathbf{x}(t), \delta \mathbf{x}'(t) \Big\rangle - \Big\langle \delta \mathbf{x}, F'^*(t)\delta \mathbf{x}'(t) \Big\rangle = 0. \end{aligned} \tag{2.20}$$

Let $\mathbf{y}$ and $\mathbf{y}'$ be any two elements of $\mathcal{E}$. Starting from the initial condition $\mathbf{y}$ at time t, the solution of the tangent linear equation (2.8) at time t' can be expressed as $R(t', t)\mathbf{y}$. Similarly, the solution of the adjoint equation (2.9), starting from $\mathbf{y}'$ at time t', can be expressed as $S(t, t')\mathbf{y}'$ at time t. From this we can write the inner products as

$$\langle R(t', t)\mathbf{y}, \mathbf{y}' \rangle = \langle \mathbf{y}, S(t, t')\mathbf{y}' \rangle, \tag{2.21}$$

which shows that $S(t, t')$ is the adjoint of $R(t', t)$. So now we can express (2.19) as

$$\nabla_{\mathbf{u}} J = \sum_{p=0}^{P} S(t_0, t_p)\nabla_{\mathbf{x}} H(t_p). \tag{2.22}$$

Next we consider the inhomogeneous adjoint equation

$$-\frac{d\delta \mathbf{x}'}{dt} = F'^*(t)\delta' \mathbf{x} + \nabla_{\mathbf{x}} H(t). \tag{2.23}$$

Using the properties of the resolvent, equations (2.15) and (2.16), the solution of (2.23) defined by the condition $\delta\mathbf{x}'(t_P) = 0$ is

$$\delta\mathbf{x}'(t) = \sum_{p=0}^{P} S(t, \tau_p)\nabla_{\mathbf{x}} H(\tau_p). \tag{2.24}$$

Comparing equations (2.22) and (2.24) shows that the gradient of the distance function $\nabla_{\mathbf{u}} J$, is equal to the solution at time t_0 of the inhomogeneous adjoint equation, $\delta\mathbf{x}'(t_0)$. In summary, the gradient $\nabla_{\mathbf{u}} J$ can be obtained, for given initial condition $\mathbf{u}$, by performing the following operations:

(i) Starting from $\mathbf{u}$ at time t_0, integrate the basic evolution equation (2.7) from t_0 to t_P. Store the computed values of $\mathbf{x}(t)$, from $t_0 \leq t \leq t_P$.

(ii) Starting from $\delta\mathbf{x}'(t_P) = 0$, integrate the inhomogeneous adjoint equation (2.23) backwards in time from t_P to t_0, the operator $F'^*(t)$ and the gradient $\nabla_{\mathbf{x}} H(t)$ being determined at each time t, from the values $\mathbf{x}(t)$ computed in the direct integration of (2.1). The final value $\delta\mathbf{x}'(t_0)$ is the required gradient $\nabla_{\mathbf{u}} J$.

2.4 Descent algorithms

Once the gradient of the distance function has been found, a descent algorithm can be used to reduce the distance function to a minimum. Successive estimates $\mathbf{u}_n$ of $\mathbf{u}_{\min}$ are obtained through descent steps of the form

$$\mathbf{u}_{n+1} = \mathbf{u}_n - \lambda_n D_n, \tag{2.25}$$

where, for each n, D_n is the descent direction determined from the successive gradients $\nabla_{\mathbf{u}} J(\mathbf{u}_n), \nabla_{\mathbf{u}} J(\mathbf{u}_{n-1}), \ldots$, and λ_n is an appropriate scalar. Three classical descent algorithms, which can be used in numerical experiments, are the steepest descent algorithm (in which $D_n = \nabla_{\mathbf{u}} J(\mathbf{u}_n)$ for any n), the conjugate gradient algorithm, and the quasi-Newton or variable metric algorithm. These are described in Gill (1982), Ralston (1965), and Conte (1981). The third is a combination of the conjugate gradient algorithm and of the very efficient, but memory-expensive, quasi-Newton algorithm (Buckley and Le Nir 1983). These descent algorithms are different in the minimization process, the descent

direction in which to perform the next step, and in the length of the descent step. Theory shows that the numerical efficiency of these three algorithms increases in the order above. The algorithm used depends on the problem considered.

The steepest descent algorithm is the basic method for finding an extremum, in this case finding the minimum. The basic idea is as follows. Given an approximation $\mathbf{u}_n$ to the $\mathbf{u}_{\min}$, look for the minimum nearest to $\mathbf{u}_n$ along the straight line through $\mathbf{u}_n$ in the direction of $-\nabla_{\mathbf{u}} J(\mathbf{u}_n)$. That is, set

$$\mathbf{u}_{n+1} = \mathbf{u}_n - \lambda_n \nabla_{\mathbf{u}} J(\mathbf{u}_n), \tag{2.26}$$

and take the next approximation to the $\mathbf{u}_{\min}$ to be $\mathbf{u}_{n+1}$.

The method of steepest descent guarantees a decrease in the function value, but the convergence may be very slow. An example of the slow convergence of the steepest descent method is the possibility of shuffling ineffectually back and forth searching for a minimum in a narrow valley.

Using the direction of the gradient seems to be a very logical thing to do. But if we choose the direction more carefully we can theoretically guarantee convergence in a finite number of steps. In this case the conjugate-gradient method is used. This method generates directions of search without storing a matrix, so is an advantage when the matrix is very large as in this case. The basic idea of this method is to use the steepest descent direction, $\nabla_{\mathbf{u}} J(\mathbf{u}_n)$ plus the pervious search directions, which are denoted by the vectors $\mathbf{p}_j$, $j = 0, \ldots, n-1$.

$$\mathbf{p}_n = -\nabla_{\mathbf{u}} J(\mathbf{u}_n) - \sum_{j=0}^{n-1} B_{nj} p_j, \tag{2.27}$$

where B_{nj} is a function of how the gradient changes. It can be shown that the conjugate gradient is not an infinite but a finite iterative method. In actual computation using the conjugate gradient method does not lead to the exact solution because of accumulated roundoff.

The most efficient of these methods is a combination of the conjugate gradient and quasi-Newton algorithms. The quasi-Newton method has a form

$$\mathbf{u}_{n+1} = \mathbf{u}_n - (f'(\mathbf{u}_n))^{-1} f(\mathbf{u}_n), \tag{2.28}$$

with the matrix $f'(u_n)$ the Jacobian matrix for f at $\mathbf{u}_n$. This algorithm is more efficient but requires much more storage, n^2 storage locations, compared to $7n$ storage location for the conjugate gradient algorithm. By combining these, an alogrithm can be obtained with good convergence properties and low storage requirements. The quasi-Newton method will be implemented first until storage becomes limited, then the conjugate gradient method will be used. Ideally, the method which combines the quasi-Newton and the conjugate-gradient method should be used.

2.5 Comments on the adjoint method

So what is an adjoint equation? It is not a backward integration of the basic dynamical equations. The quantities produced at a given time t by an adjoint integration are not physical fields at time t, but partial derivatives of the function J with respect to the initial physical fields. The difference between an adjoint integration and a backward integration of the basic dynamical equations becomes significant when the basic equations contain diffusive or dissipation terms, whose backward integration is usually, from a mathematical point of view, an ill-posed problem. Wherever the integration of the basic equation is well-posed only for integration into the future, the integration of the corresponding adjoint equation will be well-posed only for integration into the past. A stability argument developed by Talagrand (1981a) shows that a sufficient condition for convergence for a forward-backward assimilation model developed by Morel (1971), is that the linearized perturbation equation be antisymmetric. In this paper's notation, this condition means $F'^*(t) = -F'(t)$. An antisymmetric equation is identical with its adjoint, so that integrating the full equation backwards in time is the same thing as integrating the adjoint. It clearly appears that using the adjoint equation is the mathematically proper and rigorous way to achieve the original goal which was heuristically assigned to forward-backward assimilation, namely adjusting a model to observations distributed in time.

Courtier and Talagrand (1987) applied the adjoint method data to assimilation and some interesting results were found. In those experiments the height and wind fields were reconstructed even in areas which are void of data, and the process was able to infer the

large-scale fields. This property of the adjoint method is particularly valuable in tropics where observations are sparse. One process through which information is propagated and fields are adjusted to the observations is advection by the flow. Since one model solution is adjusted over the model area to the data available over a period of time, there is not only downstream advection into the future as in one-way assimilation procedures (Ghil, 1981), but also upstream advection into the past.

In Courtier and Talagrand's study of the variations of the model fields in the course of the minimization process shows that, when starting from a state of rest, the first descent step reconstructs the latitudinal gradient of geopotential and the corresponding zonal wind. In the following ten or so steps, structures are progressively built up over data-rich areas and their immediate vicinity with no associated modification of the fields elsewhere. From that time on, continuation of the minimization process modifies the fields only over data-poor areas, with no further significant decrease of the distance function. Their study also showed that small-scale noise developed especially in the data-poor areas. What happens is that the minimization process (marginally) decreases the distance function by creating unrealistic noise, especially in data-poor areas, where this noise can develop freely. Some of their results show that, everything else being the same, the amount of noise is reduced when the length of the assimilation period is increased. It is probable that the noise could be reduced to an acceptable level if observations were more numerous in space and time (Cox, 1984), which could be a problem in our case, since there is little data available in the tropics.

One possible way of avoiding or reducing the small-scale noise could be to interrupt the minimization before noise starts appearing. Another possibility is to add a penalty term measuring the amount of small-scale noise in the fields to the distance function. The presence of such a term will limit the amount of noise in the adjusted fields. In Courtier and Talagrand's experiments this additional term was

$$\mathcal{L} = A \sum^{N} \frac{1}{n+1} |\zeta_{\min}|^2, \tag{2.29}$$

where A is equal to $5 \times 10^{-4} \mathrm{m}^2$ and the summation extends over all harmonics resolved by the model. This equation is added to the distance function and reduces the amount of

noise. This may not be the best method to avoid unrealistic noise, but it is clear that the addition of a penalty term is an efficient way of limiting the amount of noise.

2.6 Adjoint method for the nondivergent barotropic model

Using this method, the derivation for the nondivergent barotropic vorticity equation proceeds as follows. First, the tangent linear equation is derived by linearizing the nondivergent barotropic vorticity equation about the present solution. Then, the adjoint of the tangent linear equation is derived. By integrating the original nonlinear model forward in time from t_0 to t_P, and then backwards in time from t_P to t_0 using the adjoint equation, we get $\nabla_{\zeta_0} J$. The gradient of the distance function, $\nabla_{\zeta_0} J$, is then used to modify the initial vorticity ζ_0 and give a better track simulation.

2.6.1 Derivation of the adjoint of the vorticity equation

The model simulates nondivergent barotropic motion on a beta plane, which is expressed through the vorticity equation

$$\frac{\partial}{\partial t}\nabla^2\psi + \frac{\partial(\psi, \nabla^2\psi + f)}{\partial(x,y)} = 0, \tag{2.30}$$

where ψ is the stream function, $\nabla^2\psi$ is the relative vorticity, and β is the meridional change of the coriolis parameter df/dy and $\partial(\ ,\ )/\partial(x,y)$ is the Jacobian. This is similar to equation (2.7). If ψ and $\psi + \delta\psi$ are two solutions to equation (2.30), the equation for $\delta\psi$ is

$$\frac{\partial}{\partial t}\nabla^2\delta\psi = \frac{\partial(\nabla^2\psi + \nabla^2\delta\psi + f, \psi + \delta\psi)}{\partial(x,y)} - \frac{\partial(\nabla^2\psi + f, \psi)}{\partial(x,y)}. \tag{2.31}$$

If we linearize about the solution ψ, equation (2.31) reduces to

$$\frac{\partial}{\partial t}\nabla^2\delta\psi = \frac{\partial(\nabla^2\psi + f, \delta\psi)}{\partial(x,y)} + \frac{\partial(\nabla^2\delta\psi, \psi)}{\partial(x,y)}. \tag{2.32}$$

This is the tangent linear equation analogous to (2.8). We now define the inner product of two vorticity fields ζ and ζ' as

$$\langle \nabla^2\psi, \nabla^2\psi' \rangle = \frac{1}{L^2}\iint \nabla\psi \cdot \nabla\psi' dx dy, \tag{2.33}$$

where the domain of integration is $(0, L)$ in both directions. Also note that the Laplacian operator is self-adjoint, i.e.

$$\langle \psi, \nabla^2 \psi' \rangle = \langle \nabla^2 \psi, \psi' \rangle. \tag{2.34}$$

The Jacobian has, for any three scalar fields a, b, c, the following useful property

$$\iint \frac{\partial(a,b)}{\partial(x,y)} c \, dxdy = \iint a \frac{\partial(b,c)}{\partial(x,y)} \, dxdy. \tag{2.35}$$

Taking the inner product of (2.32) with $\nabla^2 \delta\psi'$ we obtain

$$\left\langle \frac{\partial}{\partial t} \nabla^2 \delta\psi, \nabla^2 \delta\psi' \right\rangle = \left\langle \frac{\partial \left(\nabla^2 \psi + f, \delta\psi\right)}{\partial(x,y)}, \nabla^2 \delta\psi' \right\rangle + \left\langle \frac{\partial \left(\nabla^2 \delta\psi, \psi\right)}{\partial(x,y)}, \nabla^2 \delta\psi' \right\rangle. \tag{2.36}$$

Let us now consider each of the terms in (2.36). The first term can be written

$$\left\langle \frac{\partial}{\partial t} \nabla^2 \delta\psi, \nabla^2 \delta\psi' \right\rangle = -\left\langle \nabla^2 \delta\psi, \frac{\partial}{\partial t} \nabla^2 \delta\psi' \right\rangle. \tag{2.37}$$

Making use of (2.35) we can rewrite the second term in (2.36) as

$$\begin{aligned} \left\langle \frac{\partial \left(\nabla^2 \psi + f, \delta\psi\right)}{\partial(x,y)}, \nabla^2 \delta\psi' \right\rangle &= \frac{1}{L^2} \iint \frac{\partial \left(\delta\psi, \nabla^2 \psi + f\right)}{\partial(x,y)} \delta\psi' \, dxdy \\ &= \frac{1}{L^2} \iint \delta\psi \frac{\partial \left(\nabla^2 \psi + f, \delta\psi'\right)}{\partial(x,y)} dxdy = \left\langle \nabla^2 \delta\psi, \frac{\partial \left(\delta\psi', \nabla^2 \psi + f\right)}{\partial(x,y)} \right\rangle. \end{aligned} \tag{2.38}$$

Again making use of (2.35) the last term in (2.36) can be rewritten as

$$\begin{aligned} \left\langle \frac{\partial \left(\nabla^2 \delta\psi, \psi\right)}{\partial(x,y)}, \nabla^2 \delta\psi' \right\rangle &= -\frac{1}{L^2} \iint \frac{\partial \left(\nabla^2 \delta\psi, \psi\right)}{\partial(x,y)} \delta\psi' \, dxdy \\ &= -\frac{1}{L^2} \iint \nabla^2 \delta\psi \frac{\partial \left(\psi, \delta\psi'\right)}{\partial(x,y)} dxdy = \left\langle \nabla^2 \delta\psi, \nabla^2 \frac{\partial \left(\delta\psi', \psi\right)}{\partial(x,y)} \right\rangle. \end{aligned} \tag{2.39}$$

Using (2.37)–(2.39) we can rewrite (2.36) as

$$\left\langle \nabla^2 \delta\psi, \frac{\partial}{\partial t} \nabla^2 \delta\psi' \right\rangle = \left\langle \nabla^2 \delta\psi, \frac{\partial \left(\nabla^2 \psi + f, \delta\psi'\right)}{\partial(x,y)} \right\rangle + \left\langle \nabla^2 \delta\psi, \nabla^2 \frac{\partial \left(\delta\psi', \psi\right)}{\partial(x,y)} \right\rangle. \tag{2.40}$$

From (2.40) we get

$$\frac{\partial}{\partial t} \nabla^2 \delta\psi' = \nabla^2 \frac{\partial(\delta\psi', \psi)}{\partial(x,y)} + \frac{\partial(\nabla^2 \psi + f, \delta\psi')}{\partial(x,y)}, \tag{2.41}$$

which is the adjoint of (2.32), and analogous with (2.9). Now, the inhomogeneous adjoint equation can be expressed as

$$\frac{\partial}{\partial t}\nabla^2\delta\psi' = \nabla^2\frac{\partial(\delta\psi',\psi)}{\partial(x,y)} + \frac{\partial(\nabla^2\psi + f,\delta\psi')}{\partial(x,y)} + \zeta(t) - \hat{\zeta}(t), \tag{2.44}$$

where contributions from $\zeta(t) - \hat{\zeta}(t)$ occur only when observation are available. The procedure to solve for $\nabla_{\zeta_0} J$ is the same as shown in section 2.3 using spatial differencing (spectral method) and time differencing (Adams-Bashforth method), which are discussed in more detail in sections 3.1 and 3.2.

Chapter 3

SPATIAL AND TIME DIFFERENCING

3.1 Spatial differencing scheme

Although most tropical models have used finite difference methods to solve the model's governing equations, a spectral method is used here. This is due to the many computational advantages over finite difference methods that spectral methods have, including much greater accuracy per degree of freedom, reduction of computational dispersion, and elimination of nonlinear instability (DeMaria, 1983, and Gottlieb and Orszag, 1977).

There are three different spectral methods known as the Galerkin, tau, and collocation methods. For each of these methods, the spatial dependence of the dependent variables is expanded in a finite series of some appropriate basis functions. The governing equations are then used to give a system of equations for the series amplitudes. The differences between these methods are the way in which boundary conditions are treated, and the way the equations for the series amplitudes are determined.

For the Galerkin method the basis functions are chosen so that they satisfy the same boundary conditions as the dependent variables, and are orthogonal with respect to some inner product. The equations for the time dependent series amplitudes are then found by substituting the series expansions into the governing equations and taking the inner product of each equation with each of the basis functions.

The tau method is a variation of the Galerkin method where the basis functions are still orthogonal, but are not required to satisfy the boundary conditions individually. Instead, extra degrees of freedom are added in such a way that the series as a whole satisfies the boundary conditions. With the exception of the extra terms in the series expansions, the time dependent series amplitudes are determined in the same way as the Galerkin method.

For the collocation or pseudospectral method, the equations for the time dependent series amplitudes are determined by substituting the series expansions into the governing equations and then forcing the equations to be satisfied exactly on a set of grid points (collocation points), where the number of collocation points is chosen to be equal to the number of terms in the series expansions. For collocation, the equations are solved in physical space, while for the Galerkin and tau methods, the equations are solved in spectral space. In this study the Galerkin method is used, because of its simplicity and accuracy. With the Galerkin method, the Fourier components are eigenfunctions of the linear operators which appear in the governing equations, so the linear terms have a very simple form in spectral space. Also, with the Galerkin method no aliasing errors are produced because the equations are solved in spectral space. The collocation method, however, produces aliasing errors since the equations are solved in physical space.

For this case the appropriate basis functions are Fourier components. This reduces the need for the use of implicit time differencing since the Fourier components have uniform resolution over the domain. It turns out that when Fourier components are used as basis functions, semi-implicit time differencing can be implemented quite easily since the Fourier components are the eigenfunctions of the linear operators which appear. It is not necessary to solve a linear system at each time step since all the series amplitudes which appear in the linear terms are decoupled.

Because of the nonlinear terms in this problem a double Fourier transform is used to transfer the nonlinear terms into physical space, calculate these terms using grid points and then transfer the terms back to spectral space.

In this section the Galerkin method is applied to the governing equations to give an approximate solution. First the nondivergent barotropic vorticity equation will be solved, followed by the adjoint equation derived in chapter 2. The nondivergent vorticity equation can be expressed as

$$\frac{\partial \zeta}{\partial t} + \frac{\partial (\psi, \zeta + f)}{\partial (x, y)} = 0, \tag{3.1}$$

which if expanded, noting that $df/dy = \beta, -\partial\psi/\partial y = u$ and $\partial\psi/\partial x = v$, where u and v are the zonal and meridional wind components, is

$$\frac{\partial\zeta}{\partial t} + \beta\frac{\partial\psi}{\partial x} = F(x,y), \tag{3.2}$$

where

$$F(x,y) = -\left[\frac{\partial(u\zeta)}{\partial x} + \frac{\partial(v\zeta)}{\partial y}\right]. \tag{3.3}$$

To transform the equation a double Fourier transform is used where the Fourier transform of an unknown function such as $\psi(x,y,t)$ is denoted as $\psi_m(y,t)$, and the transform of $\psi_m(y,t)$ is denoted as $\psi_{m,n}(t)$. These transforms can be written as

$$\psi_m(y,t) = \frac{1}{L}\int_0^L \psi(x,y,t)e^{-ikx}dx, \tag{3.4}$$

$$\psi_{m,n}(t) = \frac{1}{L}\int_0^L \psi_m(y,t)e^{-ily}dy, \tag{3.5}$$

where $k = 2\pi m/L$ and $l = 2\pi n/L$. The truncated inverses of these transforms are

$$\psi_m(y,t) = \sum_{n=-N}^{N} \psi_{m,n}(t)e^{ily}, \tag{3.6}$$

$$\psi(x,y,t) = \sum_{m=-M}^{M} \psi_m(y,t)e^{ikx}. \tag{3.7}$$

Equations (3.4) and (3.5) take a function in physical space, $\psi(x,y,t)$ and transform it to spectral space, $\psi_{m,n}(t)$. Equations (3.6) and (3.7) transform a function in spectral space back into physical space.

Before proceeding here are some useful operational properties

$$\frac{1}{L}\int_0^L \frac{\partial\psi(x,y,t)}{\partial x}e^{-ikx}dx = ik\psi_m(y,t), \tag{3.8}$$

$$\frac{1}{L}\int_0^L \frac{\partial\psi_m(y,t)}{\partial y}e^{-ily}dy = il\psi_{m,n}(t). \tag{3.9}$$

Letting $u\zeta = A$ and $v\zeta = B$, (3.2) can be expressed as

$$\frac{\partial\zeta}{\partial t} + \beta\frac{\partial\psi}{\partial x} = -\left[\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y}\right]. \tag{3.10}$$

The transform of (3.10) in the x-direction is

$$\frac{\partial\zeta_m}{\partial t} + \beta ik\psi_m = -\left[ikA_m + \frac{\partial B_m}{\partial y}\right]. \tag{3.11}$$

Now transform (3.11) in the y-direction

$$\frac{\partial \zeta_{m,n}}{\partial t} + \beta ik\psi_{m,n} = -\left[ikA_{m,n} + ilB_{m,n}\right]. \tag{3.12}$$

Using the double Fourier transform, $\zeta_{m,n}$, $u_{m,n}$, and $v_{m,n}$ are the following

$$\zeta_{m,n} = -(k^2 + l^2)\psi_{m,n}, \tag{3.13}$$

$$u_{m,n} = -il\psi_{m,n}, \tag{3.14}$$

$$v_{m,n} = ik\psi_{m,n}. \tag{3.15}$$

Plugging these into equation (3.12) and rearranging, we get

$$\frac{\partial \psi_{m,n}}{\partial t} - \left(\frac{ik\beta}{k^2 + l^2}\right)\psi_{m,n} = -\frac{1}{k^2 + l^2}F_{m,n}, \tag{3.16}$$

where

$$F_{m,n} = -\left[ikA_{m,n} + ilB_{m,n}\right]. \tag{3.17}$$

To solve (3.16) the following steps are performed.

1. Compute the wind components and relative vorticity at points on the transform grid:

$$u(x_i, y_j) = -\sum_{m=-M}^{M} e^{ikx_i}\left[\sum_{n=-N}^{N} il\psi_{m,n}e^{ily_j}\right], \tag{3.18}$$

$$v(x_i, y_j) = \sum_{m=-M}^{M} e^{ikx_i}\left[\sum_{n=-N}^{N} ik\psi_{m,n}e^{ily_j}\right], \tag{3.19}$$

$$\zeta(x_i, y_j) = -\sum_{m=-M}^{M} e^{ikx_i}\left[\sum_{n=-N}^{N} \left(k^2 + l^2\right)\psi_{m,n}e^{ily_j}\right]. \tag{3.20}$$

2. Compute the relative vorticity fluxes at points on the transform grid and then do a Fourier transform in the x and y directions:

$$A(x_i, y_j) = u(x_i, y_j)\zeta(x_i, y_j), \tag{3.21}$$

$$B(x_i, y_j) = v(x_i, y_j)\zeta(x_i, y_j). \tag{3.22}$$

Now using these flux equations (3.21) and (3.22)

$$A_{m,n} = \frac{1}{(3M+1)(3N+1)} \sum_{i=1}^{3M+1} \sum_{j=1}^{3N+1} A(x_i, y_j) e^{-i(kx_i + ly_j)}, \tag{3.23}$$

$$B_{m,n} = \frac{1}{(3M+1)(3N+1)} \sum_{i=1}^{3M+1} \sum_{j=1}^{3N+1} B(x_i, y_j) e^{-i(kx_i + ly_j)}, \tag{3.24}$$

where a $3M+1$ and $3N+1$ point trapezoidal quadrature is used in the transformed space. The reason for this is it can be shown that at least this many points are needed to insure that the model is free of aliasing error.

3. Compute $F_{m,n}$, where $F_{m,n} = -[ikA_{m,n} + ilB_{m,n}]$.

The adjoint equation discussed in chapter 2 can be written as

$$\frac{\partial}{\partial t}(\delta\zeta') = \nabla^2 \frac{\partial(\delta\psi', \psi)}{\partial(x,y)} + \frac{\partial(\zeta + f, \delta\psi')}{\partial(x,y)}. \tag{3.25}$$

Expanding the first term on the right side of (3.2), and using $-\partial\psi/\partial y = u$ and $\partial\psi/\partial x = v$, we obtain

$$\nabla^2 \frac{\partial(\delta\psi', \psi)}{\partial(x,y)} = -\nabla^2 \left[\frac{\partial}{\partial x}(u\delta\psi') + \frac{\partial}{\partial y}(v\delta\psi') \right]. \tag{3.26}$$

Expanding the last term on the right side of (3.25), and noting that $-\partial\delta\psi'/\partial y = \delta u', \partial\delta\psi'/\partial x = \delta v'$ and $df/dy = \beta$, we obtain

$$\frac{\partial(\zeta + f, \delta\psi')}{\partial(x,y)} = -\left[\frac{\partial}{\partial x}(\delta u'\zeta) + \frac{\partial}{\partial y}(\delta v'\zeta) \right] - \beta \frac{\partial\delta\psi'}{\partial x}. \tag{3.27}$$

Now, let $u\delta\psi' = C, v\delta\psi' = D, \delta u'\zeta = G$ and $\delta v'\zeta = H$, and we get a simplified equation similar to (3.10)

$$\frac{\partial}{\partial t}(\delta\zeta') + \beta \frac{\partial\delta\psi'}{\partial x} = -\left[\frac{\partial}{\partial x}(\nabla^2 C + G) + \frac{\partial}{\partial y}(\nabla^2 D + H) \right]. \tag{3.28}$$

From here we can solve (3.28) in the same fashion as the nondivergent vorticity equation. First, transform (3.28) in the x-direction

$$\frac{\partial}{\partial t}(\delta\zeta'_m) + \beta ik\delta\psi'_m = -\left[ik(\nabla^2 C_m + G_m) + \frac{\partial}{\partial y}(\nabla^2 D_m + H_m) \right]. \tag{3.29}$$

Now transform in the y-direction

$$\frac{\partial}{\partial t}(\delta\zeta'_{m,n}) + \beta ik\delta\psi'_{m,n} = -\left[ik(\nabla^2 C_{m,n} + G_{m,n}) + il(\nabla^2 D_{m,n} + H_{m,n}) \right]. \tag{3.30}$$

Along with equations (3.13)–(3.15), $\delta\zeta'_{m,n}, \delta u'_{m,n}, \delta v'_{m,n}$ and ∇^2 need to be found using a double Fourier transform

$$\delta\zeta'_{m,n} = -(k^2 + l^2)\delta\psi'_{m,n}, \tag{3.31}$$

$$\delta u'_{m,n} = -il\delta\psi'_{m,n}, \tag{3.32}$$

$$\delta v'_{m,n} = ik\delta\psi'_{m,n}, \tag{3.33}$$

$$\nabla^2 = -(k^2 + l^2). \tag{3.34}$$

Using these equations, (3.30) can be rewritten as

$$\frac{\partial}{\partial t}(\delta\psi'_{m,n}) - \left(\frac{ik\beta}{k^2 + l^2}\right)\delta\psi'_{m,n} = \frac{1}{k^2 + l^2}E_{m,n}, \tag{3.35}$$

where

$$E_{m,n} = ik[-(k^2 + l^2)C_{m,n} + G_{m,n}] + il[-(k^2 + l^2)D_{m,n} + H_{m,n}] \tag{3.36}$$

As before, the terms in (3.36) are transformed into physical space, calcu ated, and then transformed back into spectral space.

1. Along with equations (3.18)–(3.20), $\delta u', \delta v'$ and $\delta\psi'$need to be computed at points on the transform grid:

$$\delta u'(x_i, y_j) = -\sum_{m=-M}^{M} e^{ikx_i}\left[\sum_{n=-N}^{N} il\delta\psi'_{m,n}e^{ily_j}\right], \tag{3.37}$$

$$\delta v'(x_i, y_j) = \sum_{m=-M}^{M} e^{ikx_i}\left[\sum_{n=-N}^{N} ik\delta\psi'_{m,n}e^{ily_j}\right], \tag{3.38}$$

$$\delta\psi'(x_i, y_j) = \sum_{m=-M}^{M} e^{ikx_i}\left[\sum_{n=-N}^{N} \delta\psi'_{m,n}e^{ily_j}\right]. \tag{3.39}$$

2. Compute the fluxes at points on the transform grid and then do a Fourier transform in the x and y directions:

$$C(x_i, y_j) = u(x_i, y_j)\delta\psi'(x_i, y_j), \tag{3.40}$$

$$D(x_i, y_j) = v(x_i, y_j)\delta\psi'(x_i, y_j), \tag{3.41}$$

$$G(x_i, y_j) = \delta u'(x_i, y_j)\zeta(x_i, y_j), \tag{3.42}$$

$$H(x_i, y_j) = \delta v'(x_i, y_j)\zeta(x_i, y_j). \tag{3.43}$$

Using these flux equations (3.40)–(3.43)

$$C_{m,n} = \frac{1}{(3M+1)(3N+1)} \sum_{i=1}^{3M+1} \sum_{j=1}^{3N+1} C(x_i, y_j) e^{-i(kx_i + ly_j)}, \tag{3.44}$$

$$D_{m,n} = \frac{1}{(3M+1)(3N+1)} \sum_{i=1}^{3M+1} \sum_{j=1}^{3N+1} D(x_i, y_j) e^{-i(kx_i + ly_j)}, \tag{3.45}$$

$$G_{m,n} = \frac{1}{(3M+1)(3N+1)} \sum_{i=1}^{3M+1} \sum_{j=1}^{3N+1} G(x_i, y_j) e^{-i(kx_i + ly_j)}, \tag{3.46}$$

$$H_{m,n} = \frac{1}{(3M+1)(3N+1)} \sum_{i=1}^{3M+1} \sum_{j=1}^{3N+1} H(x_i, y_j) e^{-i(kx_i + ly_j)}. \tag{3.47}$$

3. Compute $E_{m,n}$, where $E_{m,n} = ik[-(k^2+l^2)C_{m,n} + G_{m,n}] + il[-(k^2+l^2)D_{m,n} + H_{m,n}]$.

The Fourier series is truncated with a maximum wavenumber of M and N in the x and y directions respectively. While in the transformed space the truncated Fourier series maximum wavenumbers are $3M$ and $3N$. From the theory of numerical quadrature (Krylov, 1962) it can be shown that $3M+1$ and $3N+1$ point trapezoidal quadratures are exact if the integrand is a truncated Fourier series with a maximum wavenumber less than or equal to $3M$ or $3N$. Thus the nonlinear terms are evaluated exactly when transformed from physical to spectral space, so no aliasing error is introduced. This prevents nonlinear instability of the type described by Phillips (1959).

To reduce the computation time the model uses Fast Fourier Transform (FFT) routines written for the Cray computer by Temperton (1983a,b,c). DeMaria (1983) showed that transforming a variable from spectral to physical and back to spectral space can be evaluated about 13 times faster when the FFT algorithms are used. Since most of the computing time is used for the transforms, the use of the FFT algorithms greatly increases the efficiency of the model.

3.2 Time differencing scheme

In choosing a time differencing scheme, the accuracy, stability, and difficulties of programming must all be considered. Schemes can be divided into one-step and multi-step methods. In one-step schemes the solution at a given time step depends only upon the single state of the system at the preceding time step. Multi-step methods require the knowledge of more than one of the previous states. The multi-step methods generally have higher order accuracy but often produce nonphysical parasitic components of the solution. The one-step methods do not produce parasitic solution components, but must use smaller time steps to match the accuracy of the multi-step methods.

Talagrand and Courtier used a forward step, followed by the leapfrog time differencing scheme. However, the leapfrog scheme is a multi-step method that will produce a computational mode that is not damped with time. To avoid this problem we are using the second order Adams-Bashforth time differencing scheme. Starting from the initial conditions $\mathbf{u} = \mathbf{x}_0$ at time t_0, numerical integration of (2.7) with the Adams-Bashforth time differencing scheme, initialized by a forward step, produces the following sequence of estimates $\mathbf{x}_p$ for $\mathbf{x}(t)$ at times $t_p = t_0 + p\Delta t$, $p = 1, 2, \ldots$:

$$\mathbf{x}_1 = \mathbf{x}_0 + \Delta t F(\mathbf{x}_0) \qquad \text{(forward step)}, \tag{3.48}$$

$$\mathbf{x}_{p+1} = \mathbf{x}_p + \frac{3}{2}\Delta t F(\mathbf{x}_p) - \frac{1}{2}\Delta t F(\mathbf{x}_{p-1}) \qquad p \geq 1. \tag{3.49}$$

The Adams-Bashforth time differencing scheme shown above is in the form used for a forward integration. However, some work needs to be done to derive the time differencing scheme for integrating backwards using the adjoint equation. So, the idea here is to find the adjoint of the Adams-Bashforth time differencing scheme and the adjoint of the forward step.

To do this the discretized analogue of the distance function (2.10) is needed,

$$J = \sum_{p=0}^{P} H_p(\mathbf{x}_p), \tag{3.50}$$

where $H_p(\mathbf{x}_p)$ is expressed in (2.11), for each time step p. The first-order variation of J (2.12) is

$$\delta J = \sum_{p=0}^{P} \langle \nabla_{\mathbf{x}} H_p, \delta \mathbf{x}_p \rangle. \tag{3.51}$$

Using equations (3.48) and (3.49),we can express the Adams-Bashforth time differencing scheme for the tangent linear equation as

$$\delta \mathbf{x}_1 = (I + \Delta t F_0') \delta \mathbf{x}_0, \tag{3.52}$$

$$\delta \mathbf{x}_{p+1} = \delta \mathbf{x}_p + \frac{3}{2} \Delta t F_p' \delta \mathbf{x}_p - \frac{1}{2} \Delta t F_{p-1}' \delta \mathbf{x}_{p-1} \quad p \geq 1, \tag{3.53}$$

where I is the identity operator of $\mathcal{E}$, a Hilbert space, and F_p' is the derivative of F with respect to $\mathbf{x}$, taken at point $\mathbf{x}_p$. Introducing, for $p \geq 1$, the following vector in $\mathcal{E}^2$

$$\delta \mathbf{y}_p = \begin{pmatrix} \delta \mathbf{x}_{p-1} \\ \delta \mathbf{x}_p \end{pmatrix}. \tag{3.54}$$

Equations (3.52) and (3.53) can be expressed as

$$\delta \mathbf{y}_1 = T_0 \delta \mathbf{x}_0, \tag{3.55}$$

$$\delta \mathbf{y}_{p+1} = T_p \delta y_p \quad p \geq 1, \tag{3.56}$$

where T_0 is the operator from $\mathcal{E}$ into $\mathcal{E}^2$ represented by the block matrix

$$T_0 = \begin{pmatrix} I \\ I + \Delta t F_0' \end{pmatrix}, \tag{3.57}$$

and T_p is, for $p \geq 1$, the operator from $\mathcal{E}^2$ into itself represented by the matrix

$$T_p = \begin{pmatrix} 0 & I \\ -\frac{1}{2} \Delta t F_{p-1}' & I + \frac{3}{2} \Delta t F_p' \end{pmatrix}. \tag{3.58}$$

We denote by $\mathcal{P}$ the projection operator from $\mathcal{E}^2$ onto $\mathcal{E}$, and is expressed in matrix form as

$$\mathcal{P} = (0 \quad I). \tag{3.59}$$

With this notation $\delta \mathbf{x}_p$ $(p \geq 1)$ is equal to

$$\delta \mathbf{x}_p = \mathcal{P} T_{p-1} \ldots T_0 \delta \mathbf{x}_0. \tag{3.60}$$

Carrying this expression into equation (3.51) and taking adjoints leads to

$$\delta J = \sum_{p=0}^{P} \langle T_0^* \dots T_{p-1}^* \mathcal{P}^* \nabla_{\mathbf{x}} H_p, \delta \mathbf{x}_0 \rangle + \langle \nabla_{\mathbf{x}} H_0, \delta \mathbf{x}_0 \rangle, \tag{3.61}$$

from this the gradient of J with respect to $\mathbf{x}_0$ is equal to

$$\nabla_{\mathbf{x}_0} J = \sum_{p=0}^{P} \langle T_0^* \dots T_{p-1}^* \mathcal{P}^* \nabla_{\mathbf{x}} H_p \rangle + \nabla_{\mathbf{x}} H_0. \tag{3.62}$$

The adjoints T_p^* T_0^* and $\mathcal{P}^*$ are represented by the transposes of the corresponding matrices (3.57), (3.58), and (3.59), i.e.,

$$T_0^* = (I \quad I + \Delta t F_0'^*), \tag{3.63}$$

$$T_p^* = \begin{pmatrix} 0 & -\frac{1}{2}\Delta t F_{p-1}'^* \\ I & I + \frac{3}{2}\Delta t F_p'^* \end{pmatrix}, \tag{3.64}$$

$$\mathcal{P}^* = \begin{pmatrix} 0 \\ I \end{pmatrix}. \tag{3.65}$$

Now, we would like to expand (3.62) into the discretized adjoint of the Adams-Bashforth time differencing scheme. To do this we need to first define the following, $V_p = 0$ and $W_p = \delta \mathbf{x}_p$. Noting that $\nabla_{\mathbf{x}} H_p = \delta \mathbf{x}_p$ and using the expressions for $\mathcal{P}^*$, V_p, and W_p, the first term on the right side of (3.62) can be represented as

$$T_0^* \dots T_{p-1}^* \mathcal{P}^* \nabla_{\mathbf{x}} H_p = T_0^* \dots T_{p-1}^* \binom{V_p}{W_p}, \tag{3.66}$$

Now, using (3.64) we obtain

$$T_0^* \dots T_{p-1}^* \mathcal{P}^* \nabla_{\mathbf{x}} H_p = T_0^* \dots T_{p-2}^* \begin{pmatrix} -\frac{1}{2}\Delta t F_{p-2}'^* W_p \\ V_p + \left(I + \frac{3}{2}\Delta t F' p_{p-1}^*\right) W_p \end{pmatrix}. \tag{3.67}$$

Letting $-\frac{1}{2}\Delta t F_{p-2}'^* W_p = V_{p-1}$ and $V_p + (I + \frac{3}{2}\Delta t F_{p-1}'^*) W_p = W_{p-1}$, (3.67) becomes

$$T_0^* \dots T_{p-1}^* \mathcal{P}^* \nabla_{\mathbf{x}} H_p = T_0^* \dots T_{p-2}^* \binom{V_{p-1}}{W_{p-1}}, \tag{3.68}$$

this is continued to the Euler step

$$T_0^* \dots T_{p-1}^* \mathcal{P}^* \nabla_{\mathbf{x}} H_p = T_0^* \binom{V_1}{W_1} = (I \quad I + \Delta t F_0'^*) \binom{V_1}{W_1} = V_1 + (I + \Delta t F_0'^*) W_1. \tag{3.69}$$

So, from this we get the following equations

$$V_{p-1} = -\frac{1}{2}\Delta t F_{p-2}'^* W_p, \tag{3.70}$$

$$W_{p-1} = V_p + (I + \frac{3}{2}\Delta t F'^{*}_{p-1})W_p, \tag{3.71}$$

from $p = M, M-1, \ldots, 2$, with M being the backward time steps. Equations (3.70) and (3.71) are the adjoint of the Adams-Bashforth time differencing scheme, with the adjoint of the Euler step represented by (3.69). Specifically, the $\mathbf{x}$ we are referring to is ζ and the gradient of the distance function is $\nabla_{\zeta_0} J$, so the equations used are the following

$$\nabla_{\zeta_0} J = \sum_{p=1}^{P} \langle T_0^* \ldots T_{p-1}^* \mathcal{P}^* \delta\zeta_p \rangle + \delta\zeta_0, \tag{3.72}$$

where

$$\delta\zeta_p = (\zeta_p - \hat{\zeta}_p) = \nabla_\zeta H_p. \tag{3.73}$$

Now letting $V_M = 0$ and $W_M = \delta\zeta_M$, where $\delta\zeta_M = \delta\zeta_{P'}$, we get

$$V_{p-1} = -\frac{1}{2}\Delta t F'^{*}_{p-2} W_p, \tag{3.74}$$

$$W_{p-1} = \delta\zeta_{p-1} + V_p + (I + \frac{3}{2}\Delta t F'^{*}_{p-1})W_p, \tag{3.75}$$

for $p = M, M-1, \ldots, 2$. Contributions from $\delta\zeta_{p-1}$ only occur when when observations are available. The Euler step is expressed as

$$W_0 = \delta\zeta_0 + V_1 + (I + \Delta t F'^{*}_0)W_1, \tag{3.76}$$

and from this we get

$$\nabla_{\zeta_0} J = W_0. \tag{3.77}$$

The gradient of the distance function is what we are looking for. A descent algorithm can be applied to $\nabla_{\zeta_0} J$ to modify the initial conditions and improve the forecast.

Chapter 4

EXPERIMENTS WITH THE NONDIVERGENT BAROTROPIC MODEL.

In this chapter, numerical results from the forward integration of the nondivergent barotropic model are shown. This is a preliminary study showing the errors in the track forecast that result with this dynamical model without the use of the adjoint method. The motivation for this is to show how important the previous track of a tropical cyclone can be in forecasting its movement, and why the ajoint method is useful. The experiments use the nondivergent barotropic model described in section 2.1, with a large scale zonal wind profile shown in Fig. 4.1. This zonal wind profile is specified as

$$u = U \sin(\frac{2\pi}{L_y}) \tag{4.1}$$

where U is the maximum large scale wind and is equal to 10ms^{-1}. The zonal wind given by (4.1) corresponds to a single sine wave in the north-south direction, with easterlies in the southern half of the domain and westerlies in the northern half of the domain. As described in section 2.1, the model is truncated at wavenumber 47 with no dispersive term present. Several experiments are discussed which indicate how useful the adjoint method is. The first study compares, a so called observed vortex track, with a model vortex track. This is done by starting the model runs at the same place with the structure of the vortices slightly different. The second experiment is designed to test the sensitivity of model track to changes in the vortex structure. The third experiment shows an example of what the track of the model might be if the adjoint method was used.

4.1 Experiment I.

The first experiment proceeds in the following way. At the initial time, the bogus vortex described by (2.4) is located at x = 2500 km, and y = 1500 km. The model is

integrated forward 12 hours, and the results are used as the observed initial conditions. The vortex at 12 hours is slightly different in structure to original bogus vortex. At this same position, $x = 2195$ km, $y = 1516$ km, another model run is started using the bogus vortex. This second run we denote as the model integration. The observed and the model runs are integrated forward 48 hours, and their results are compared. The results and interpretation of these results are discussed in the next section.

4.1.1 Results

Starting at the initial conditions, we compare the observed and model vorticity and wind fields at the initial conditions in Figs. 4.2-4.5. These figures show no significant difference between the observed and model conditions. The maximum normalized vorticity for the observed field is 26.9, compared to 27.1 for the model field, with the southern part of the observed vortex slightly distorted. While there appears to be no difference between the initial observed and model wind fields. Looking more closely at the model and observed vortices, Fig. 4.6 shows the average vorticity at different distances from the center of the vortex, for both the observed and model cases at the initial time. The only noticeable difference in Fig. 4.6 is at a radius of 20 km or smaller. Looking at the percentage difference of vorticity between the two vortices, Fig 4.7, we see that the difference is around 1.0% except at the larger radii, where the difference is a maximum of 10.3%. where the positive values indicate that the model's vorticity is larger then the observed vorticity. The larger percentage differences at the larger radii can be misleading. Since the values of the vorticity get small as the radius get larger, small changes in the vorticity produce large percentage differences. The small differences in the initial conditions result in a hurricane track difference shown in Fig. 4.8. This figure shows that the model track does not curve as far northward as the observed track, and with a slightly faster average speed, 6.31ms^{-1} compared to 6.09ms^{-1}. The mean forecast error, Fig. 4.9, shows a linear relationship with time after, with an error of 159 km at 48 hours.

What is the cause of the error in the track forecast? If we look at the 48 hour observed and model vorticity and wind fields, Figs. 4.10-4.13., there is little difference in the wind fields, however the vorticity fields show a great deal of difference in the normalized

maximum vorticity. Looking at the observed and model radial profiles and differences in these profiles, Figs. 4.14 and 4.15, it is evident that Figs. 4.10 and 4.11 do not show a realistic comparison. This is because the figures use grid point values to produce the vorticity fields, and the maximum values of vorticity could be located between these grid points. The actual difference between the two vortices is only noticeable at radial distances further than 100 km, Figs. 4.14 and 4.15. Again the percentage difference at the larger radii reflect the small values of the vorticity. However, by 48 hours the difference at radii greater than 100 km is more significant, but the structure of the two vortices is still fairly close. These results are expected, since with the nondivergent barotropic model there is no development or dissipation of vorticity, only advection. The small differences between the initial vortices indicates that subtle changes in vortex and vorticity field produce significant changes in the track. This indicates that, the track itself contains important information on the vortex. This suggests that the adjoint method can be very useful in forecasting hurricane motion, since by this method, the model track is modified to fit the observed track. So with the adjoint method, the information contained in the observed track is used to forecast the hurricane motion.

4.2 Experiment II.

In this experiment we test the sensitivity of the vortex track to changes in the vortex structure. Again the initial vortex used in all the runs is the one described in section (2.1). We compared the model track in experiment I. and two other model tracks started at different times, with the observed track in experiment I. Again we start the observed run at $x = 2500$ km and $y = 1500$ km, and run the model for 66 hours. We use the results from the first experiment, where the model run started at $x = 2195$, $y = 1516$, which is the position of the observed vortex at 12 hours. Another model run was started at the position the observed vortex is located at 6 hours. This run represents a model vortex whose structure is closer to the observed vortex than the model vortex in experiment I. The third model run started at the position where the observed vortex is located at 18 hours, and represents a vortex whose structure is not as close to the observed vortex as the model vortex in experiment I, see Fig 4.16.

4.2.1 Results

The comparison of the observed and model vorticity fields at 6 hours is shown in Figs. 4.17 and 4.18. Again, with the exception of some distortion in the observed vortex, the figures look similar. There is little difference in the radial profiles, and in the difference in the average radial vorticity, Figs. 4.19 and 4.20. In Fig 4.20, as before, the largest percentage differences are away from the hurricanes inner core. These differences are, in general, less than in the first experiment, with the largest value being 6.1% at 180 km. The 18 hour observed and model vorticity fields, Figs. 4.21 and 4.22, have the same maximum value, but the observed vortex is more distorted than the observed vortices for the 6 and 12 hours. The radial profiles and differences in radial vorticity, Figs. 4.23 and 4.24, are similar to the other runs, with the maximum value of the difference in vorticity being 13.5%. We compare the difference in the average radial vorticity for all three runs, Fig. 4.25. Here we see that the structure of the vortex for the model run started at 6 hours is closer to observed vortex, followed by the 12 hour run, and finally the 18 hour run. So, how does this difference in affect the hurricane track? We would assume that the vortex track of the 6 hour run would be closer to the observed track. While the track of the 18 hour run would have the largest mean forecast error, and this is true. Figs. 4.26, 4.27, and 4.28 show the 6 hour track, the 18 hour track, and 6, 12, and 18 mean forecast error respectively. The mean forecast error track error of these tracks, Fig 4.28, shows that, as the differences between the observed and model vortices become larger, the track's mean forecast error becomes larger, with the mean forecast error at 48 hours for the 6, 12, and 18 hours being 80.5, 159.3, and 258.2 km, respectively. This seems obvious, but recall that the percentage difference between the observed and model vortices did not change much for these three cases. Also these changes in difference were mainly at radii outside the inner core of the hurricane, where the values of the vorticity is very small. So, we can conclude that small changes in the structure of the vortex at larger radii, produce significant changes in the hurricane track. This means, to force the model track to be the same as the observed track, the adjoint method would need to make subtle changes to the outer regions of the vortex.

4.3 Experiment III.

To give an indication of how much the adjoint method will improve the track forecast, we do the following experiment. To simulate the track that the adjoint model would have, we use the 48 hour vortex tracks from the first experiment and make some assumptions. The observed and model track are the ones shown in Fig. 4.8. Remember that using the adjoint method forces the model track to follow the observed track from time t_0 to t_1, where t_1 is the initial time shown in Fig. 4.8, and t_0 is, lets say, t_1 - 12 hours. Now we assume that the model using the adjoint method will follow the observed track for a certain amount of time after t_0 before starting to deviate. This is a reasonable assumption, since the vortex track of the adjoint model at t_1 is going the same direction and speed as the observed vortex. We also assume that the difference between the vortex of the adjoint model and the observed model at the time of deviation, t_2, is similar to the difference between the model and observed vortices at the initial time t_0. So, we assume that the mean forecast error with respect to time is the same for the model and adjoint model. In other words, the mean forecast error for the model at 12 hours after the initial time t_1, is equal to the mean forecast error of the adjoint model at 12 hours after t_2. Again this assumption seems reasonable since in order to force the adjoint model track to be the same as the observed track, the vortex of the adjoint model is modified. And, in order to have similar tracks, we would expect the observed and adjoint vortices to be similar. This modification should make the track of the adjoint vortex to be more like the observed track

4.3.1 Results

We run two cases to show the possible improvement of the mean forecast error by using the adjoint method. In the first case, the adjoint model track is the same as the observed track for 6 hours after the initial time t_1. Figure 4.29 compares the vortex track of the adjoint model with the tracks of the model and observed vortices, for this case. The mean forecast error for the model and adjoint model is shown in Fig. 4.30. The mean forecast error is approximately 25 km less throughout the time period when the

adjoint model track follows the observed track for 6 hours. In the second case the adjoint model track does not deviate from the observed track until 12 hours after the initial time t_1. The vortex tracks and the mean forecast errors are shown in Figs. 4.31 and 4.32, respectively. In this case the improvement in the mean forecast error by the adjoint model is, as expected, better than the first case. By 48 hours the mean forecast error for the adjoint model is 107 km compared to 159 km for the model, which is about a 33% reduction in the mean forecast error. Although these results do not use the adjoint model, they do give us an indication of how much the adjoint method will improve hurricane track forecasts. And from these results and the results in the first two experiments, we can see that the use of the adjoint method in hurricane track forecasts should be studied in more detail.

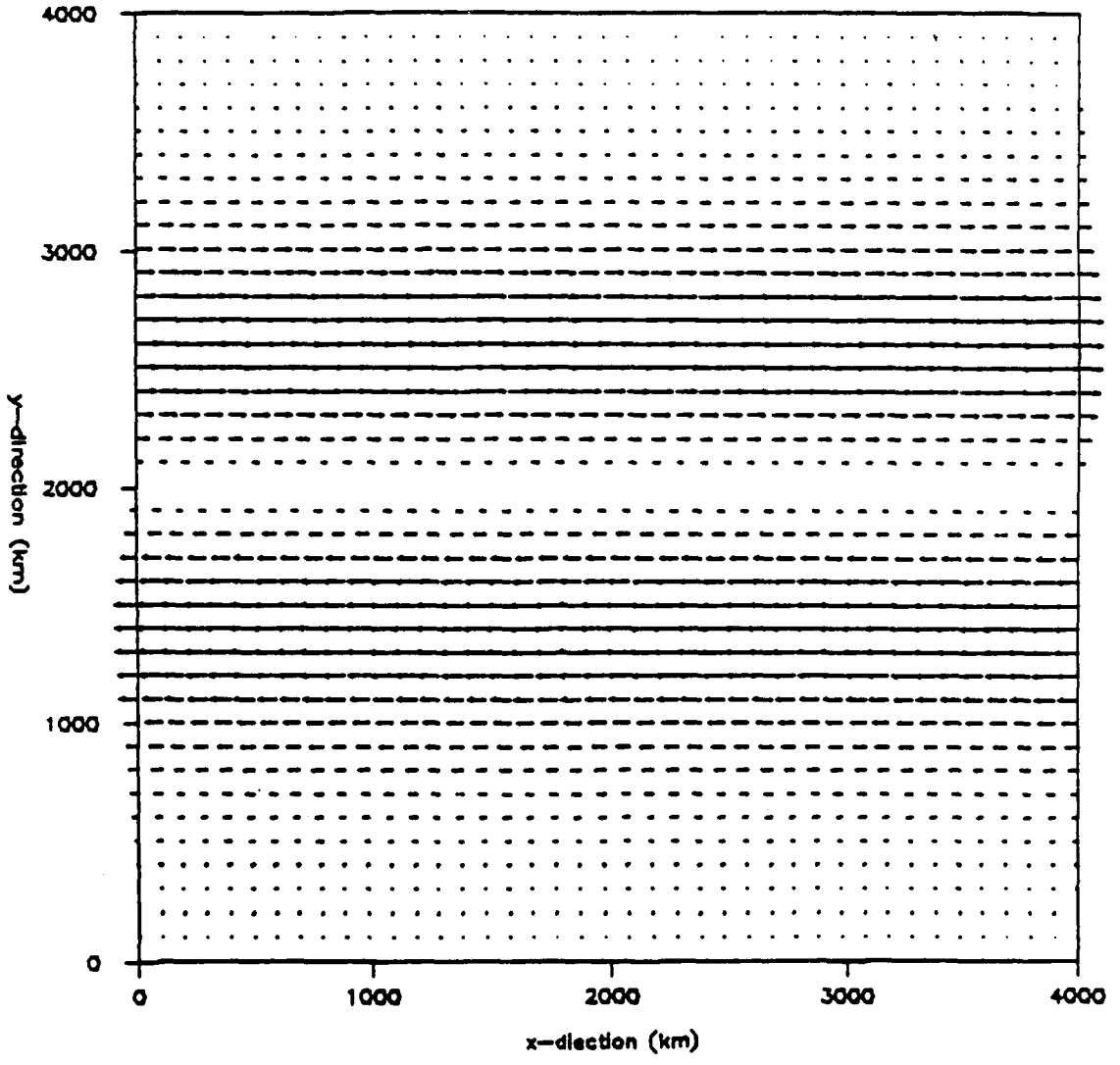


Figure 4.1: The large scale wind profile, with a maximum wind of 10 ms. The wind profile is specified by the equation $u = -10\sin(\frac{2\pi y}{L_y})$, and corresponds to a single sine wave in the north-south direction, with easterlies in the southern half of the domain and westerlies in the northern half of the domain.

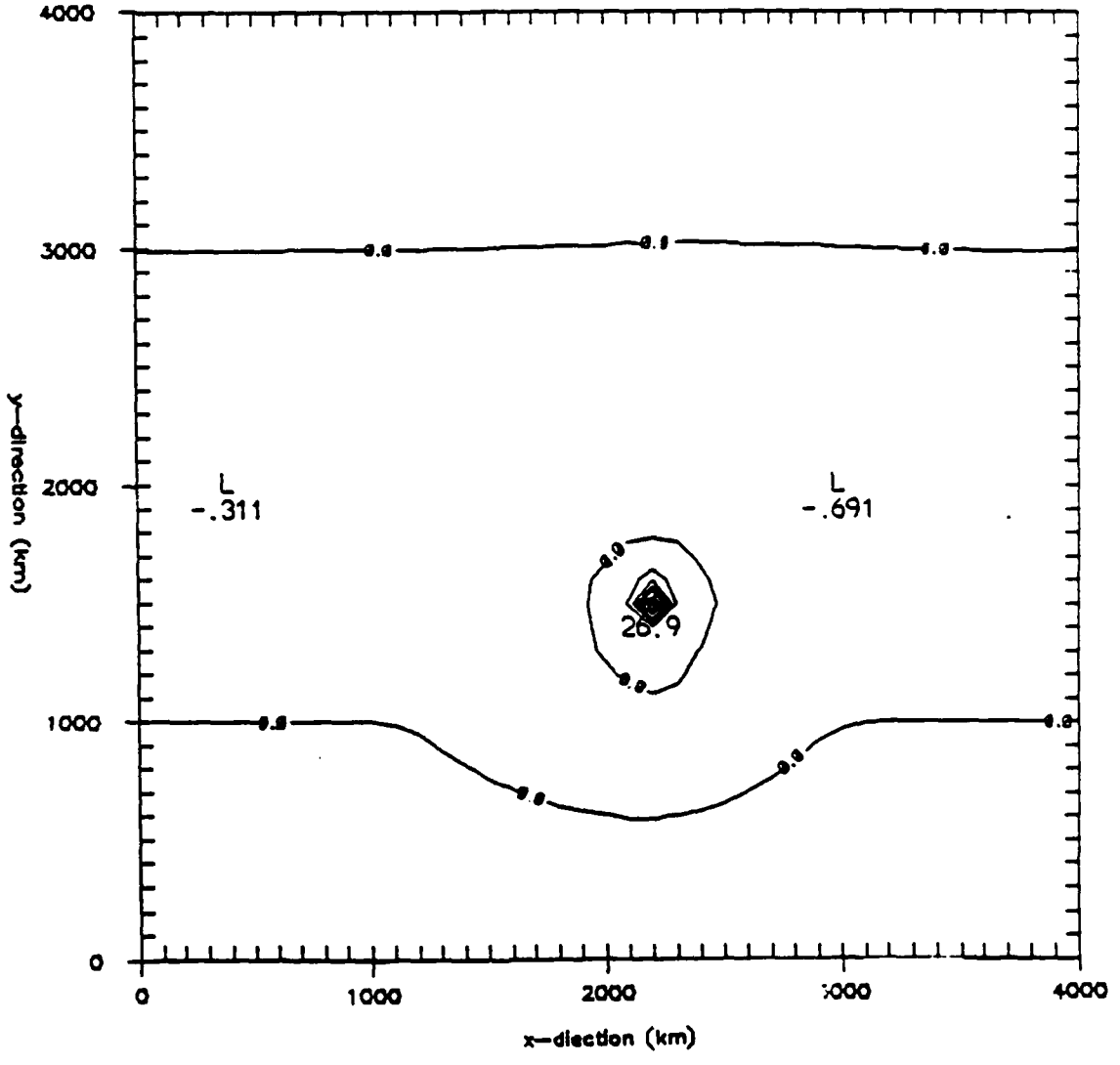


Figure 4.2: The normalized vorticity field at 12 hours for the model run started at x = 2500 km, y = 1500 km. We denote this as the initial observed vorticity field, with the vortex center at x = 2195 km, y = 1516 km.

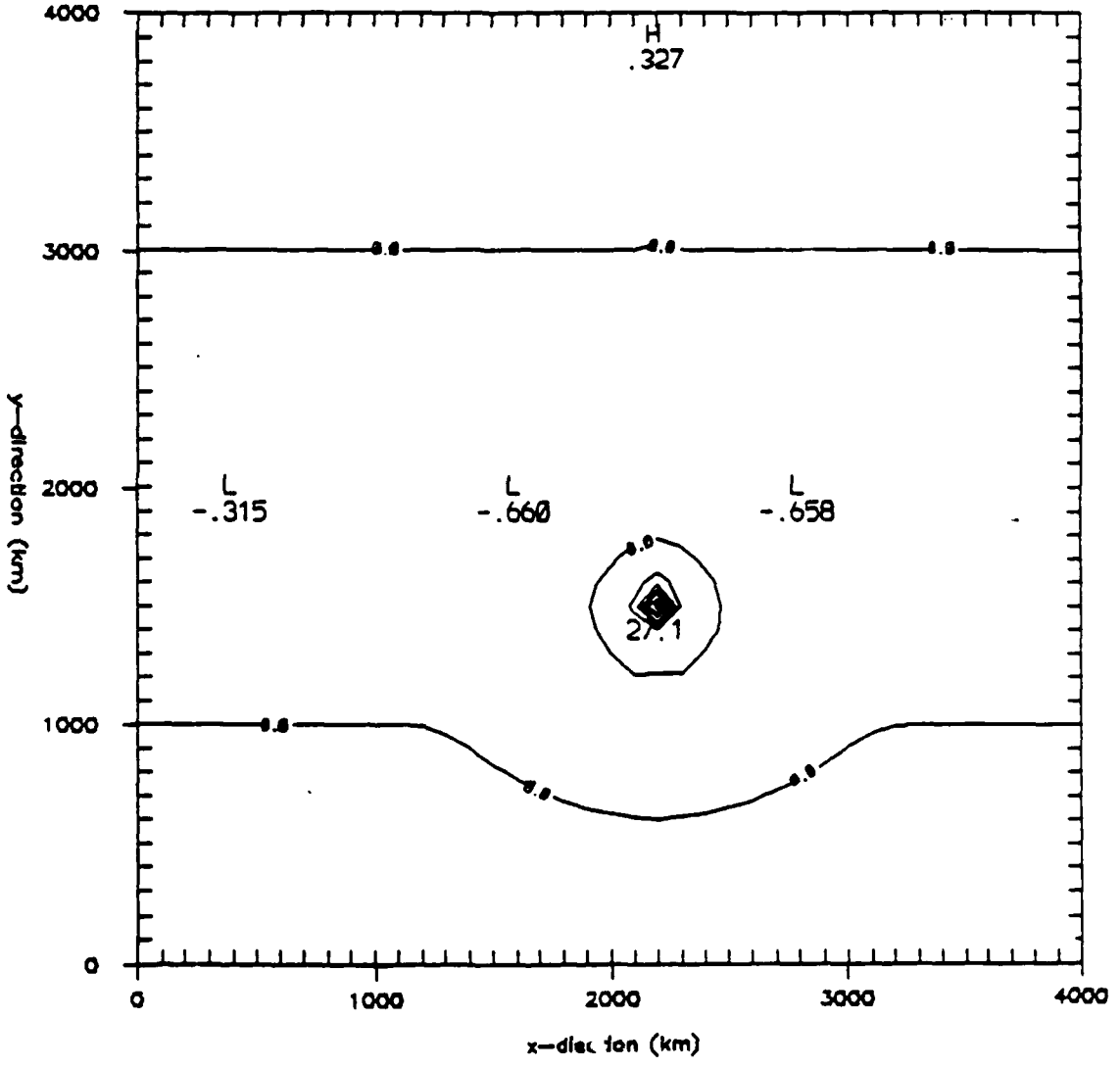


Figure 4.3: The initial normalized vorticity field for the model run, with the vortex centered at $x = 2195$ km, $y = 1516$ km.

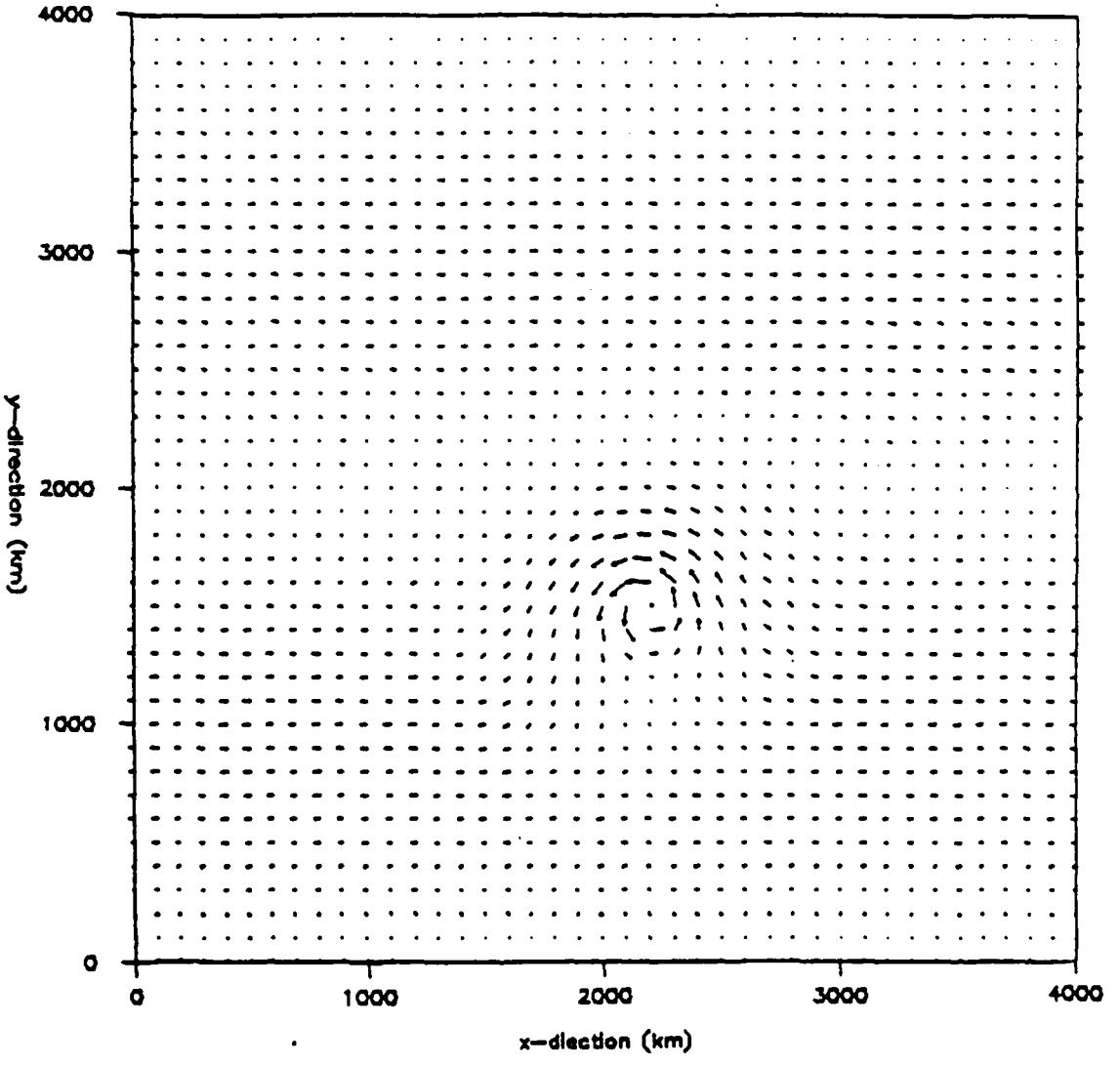


Figure 4.4: The wind field at 12 hours for the model run started at x = 2500 km, y = 1500 km, denoted as the initial observed wind field. This corresponds to Fig. 4.2.

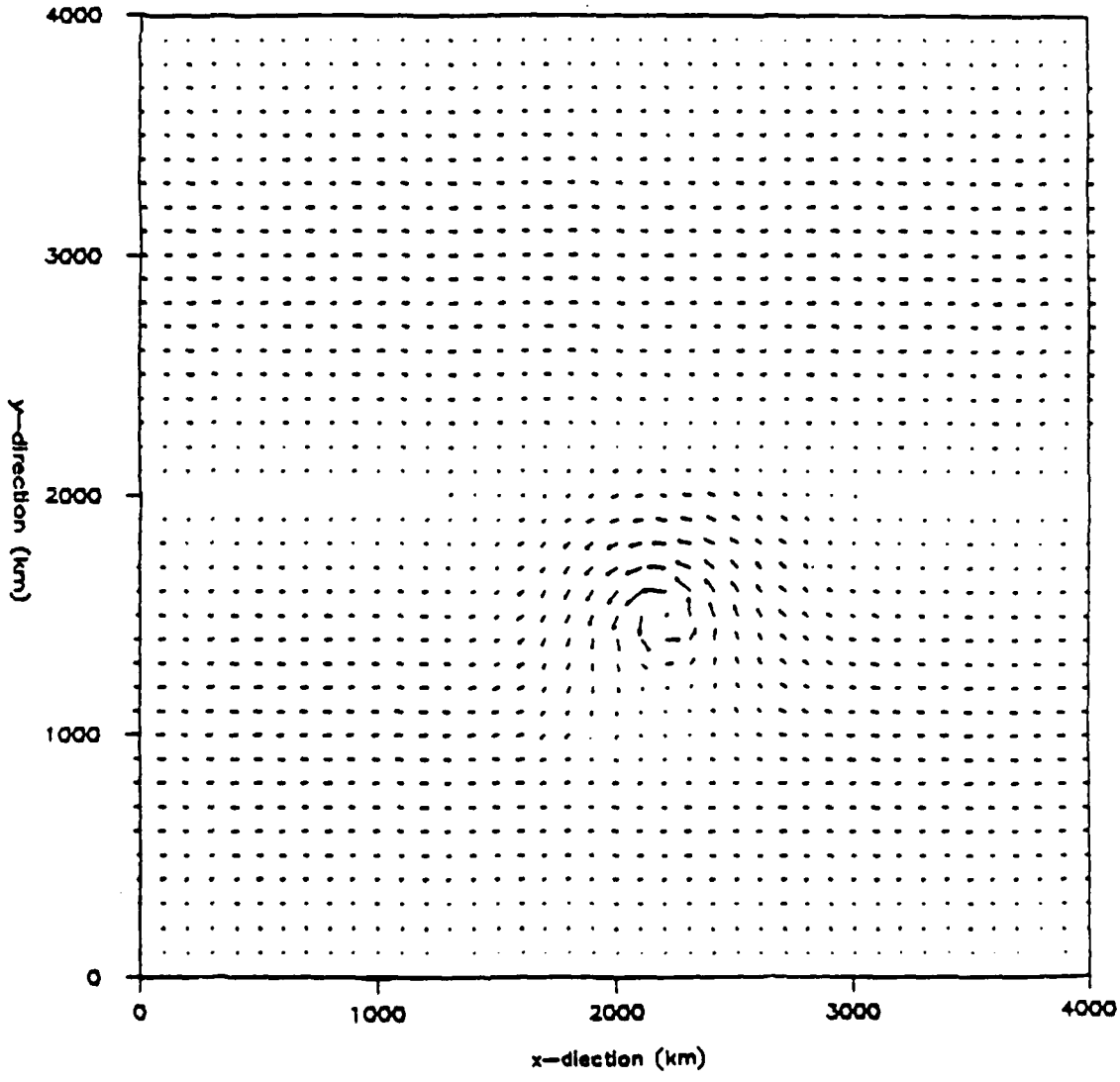


Figure 4.5: The initial wind field for the model run, corresponding to Fig. 4.3.

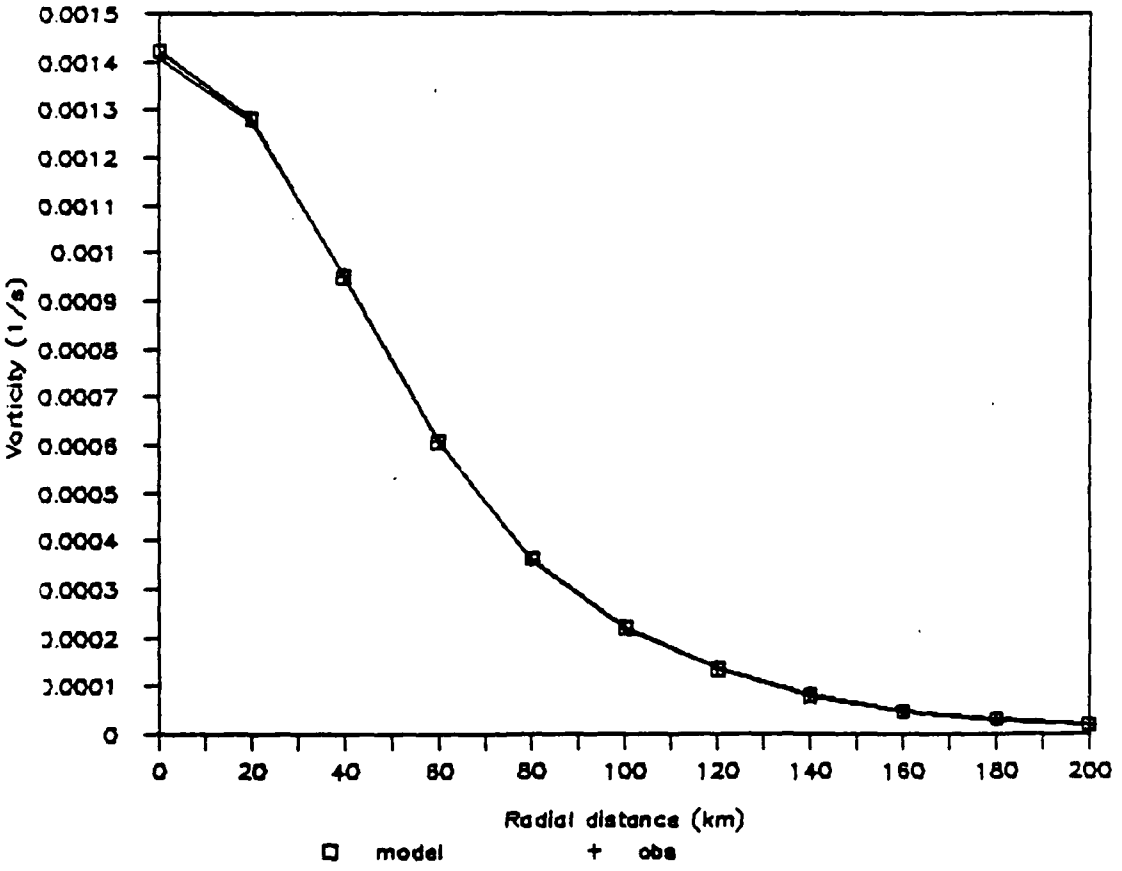


Figure 4.6: The radial profile of the vortices at the initial time. This figure compares average radial vorticity of the observed and model vortices.

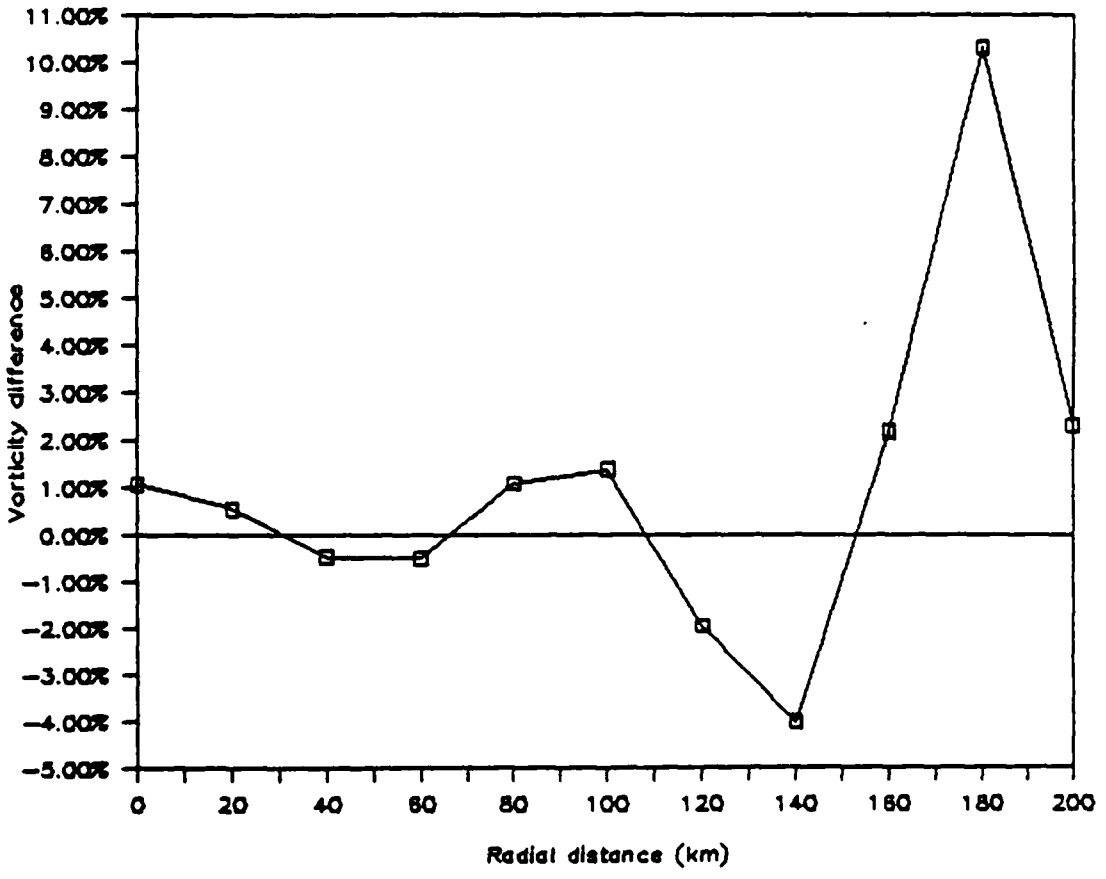


Figure 4.7: The precentage difference in the average radial vorticity between observed and model initial vortices.

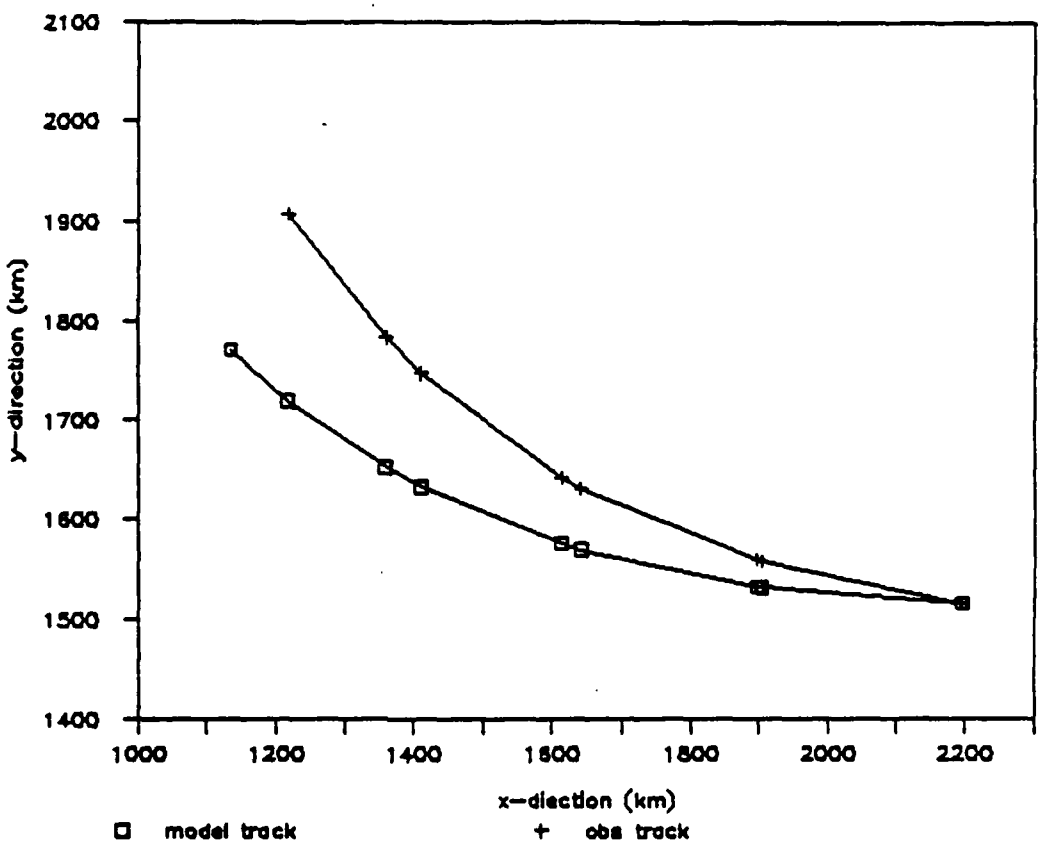


Figure 4.8: The observed and model 48 hour vortex tracks.

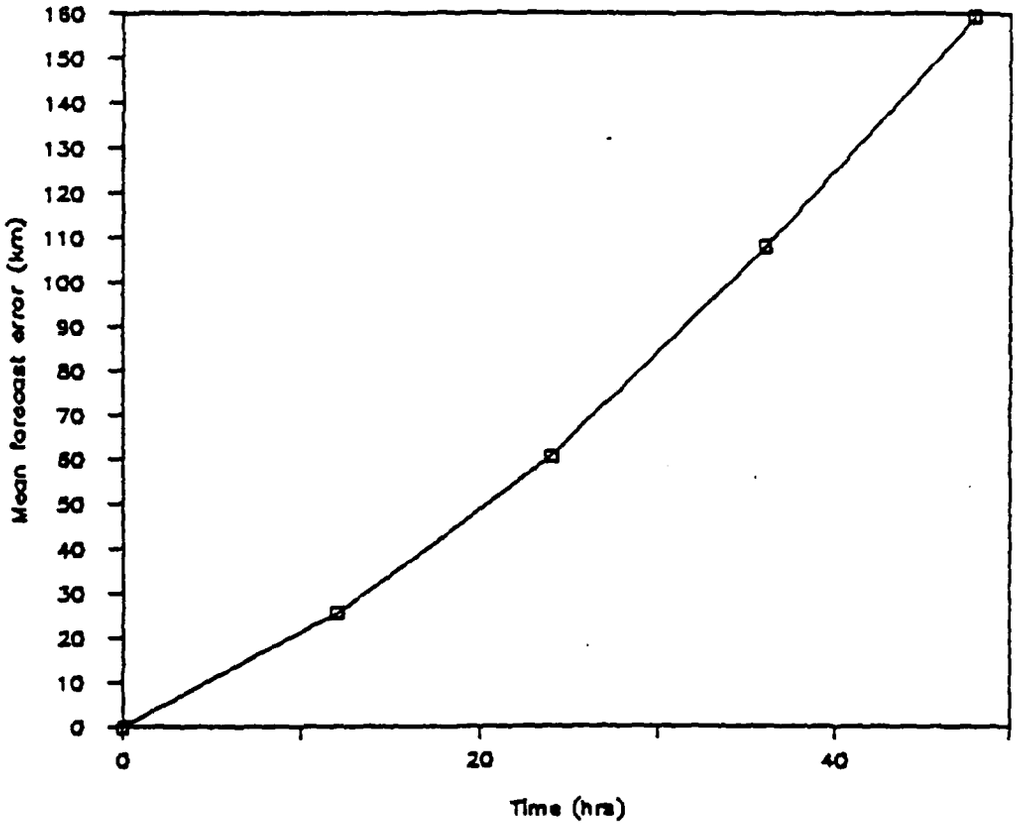


Figure 4.9: The mean forecast error of the model vortex track from 0 to 48 hours.

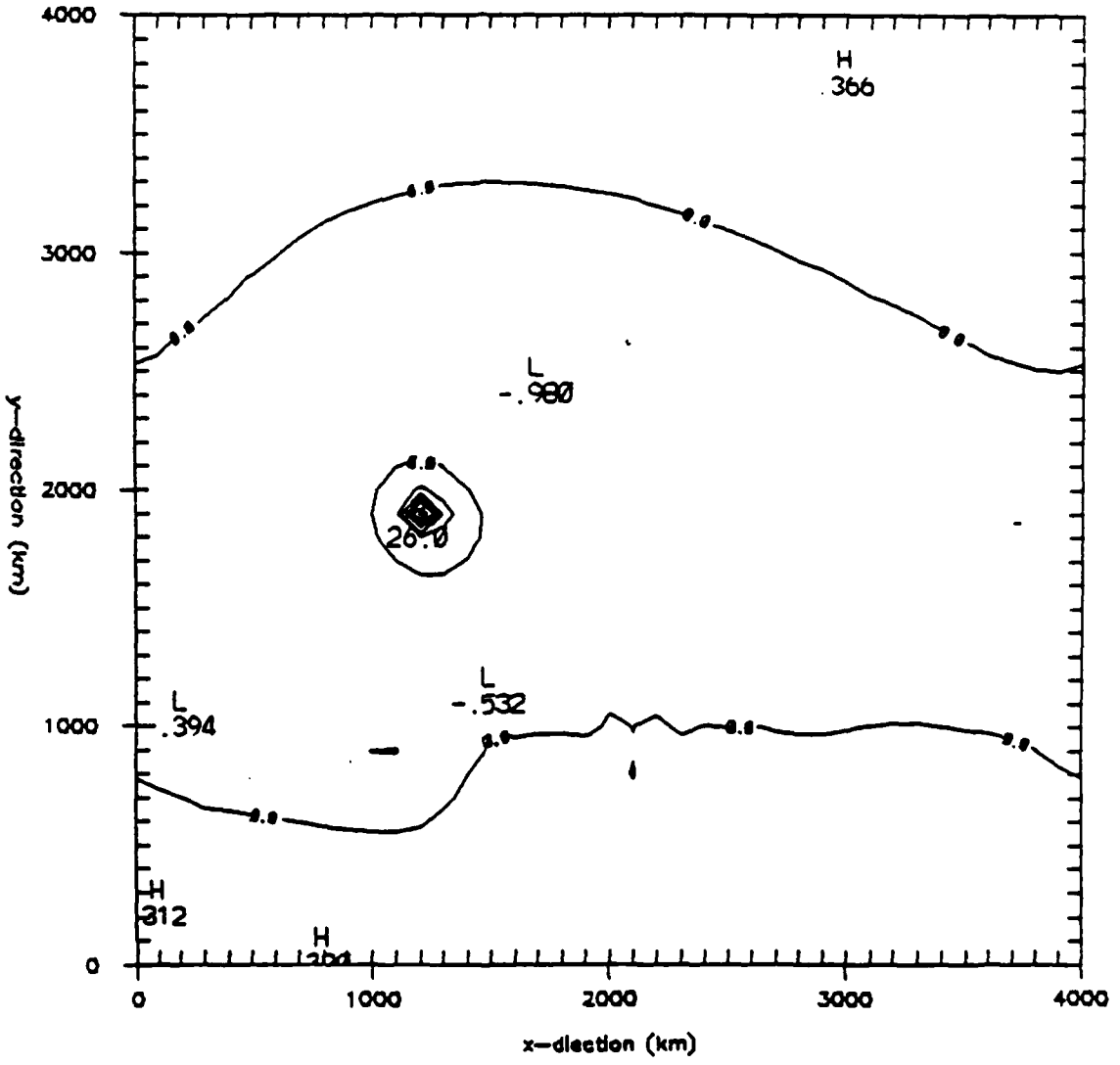


Figure 4.10: The observed normalized vorticity field at 48 hours.

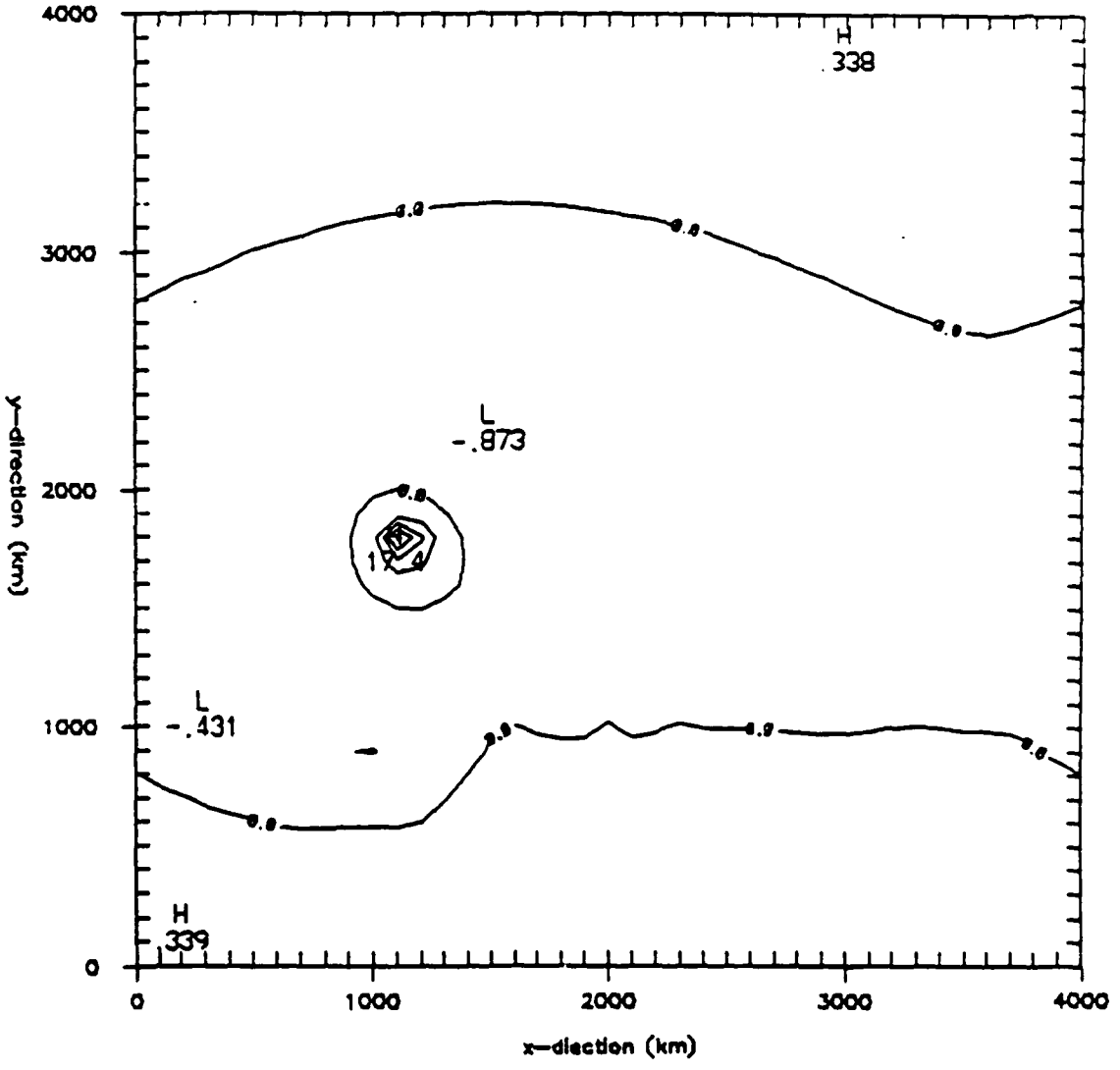


Figure 4.11: The model normalized vorticity field at 48 hours.

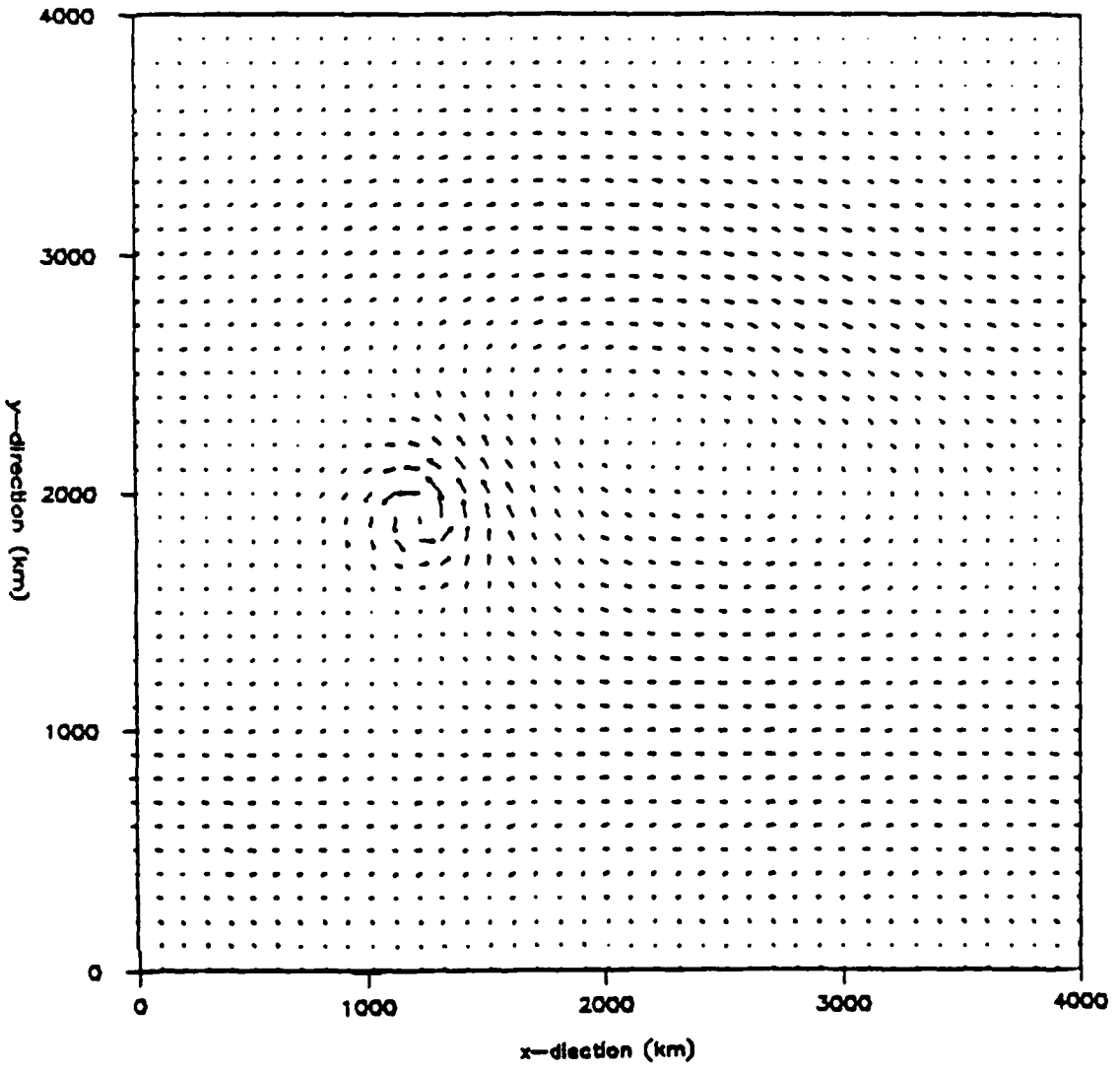


Figure 4.12: The observed wind field at 48 hours.

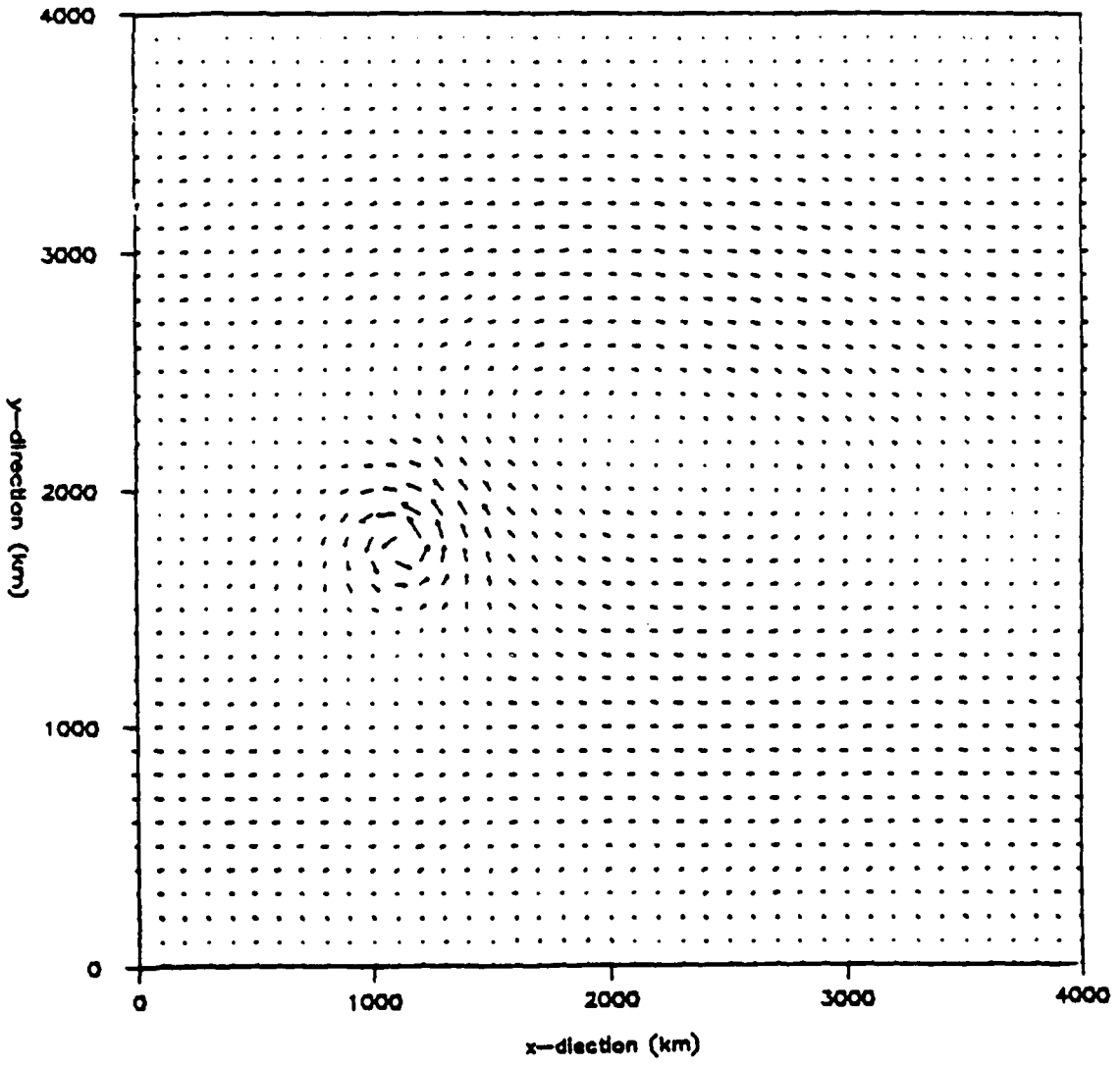


Figure 4.13: The model wind field at 48 hours.

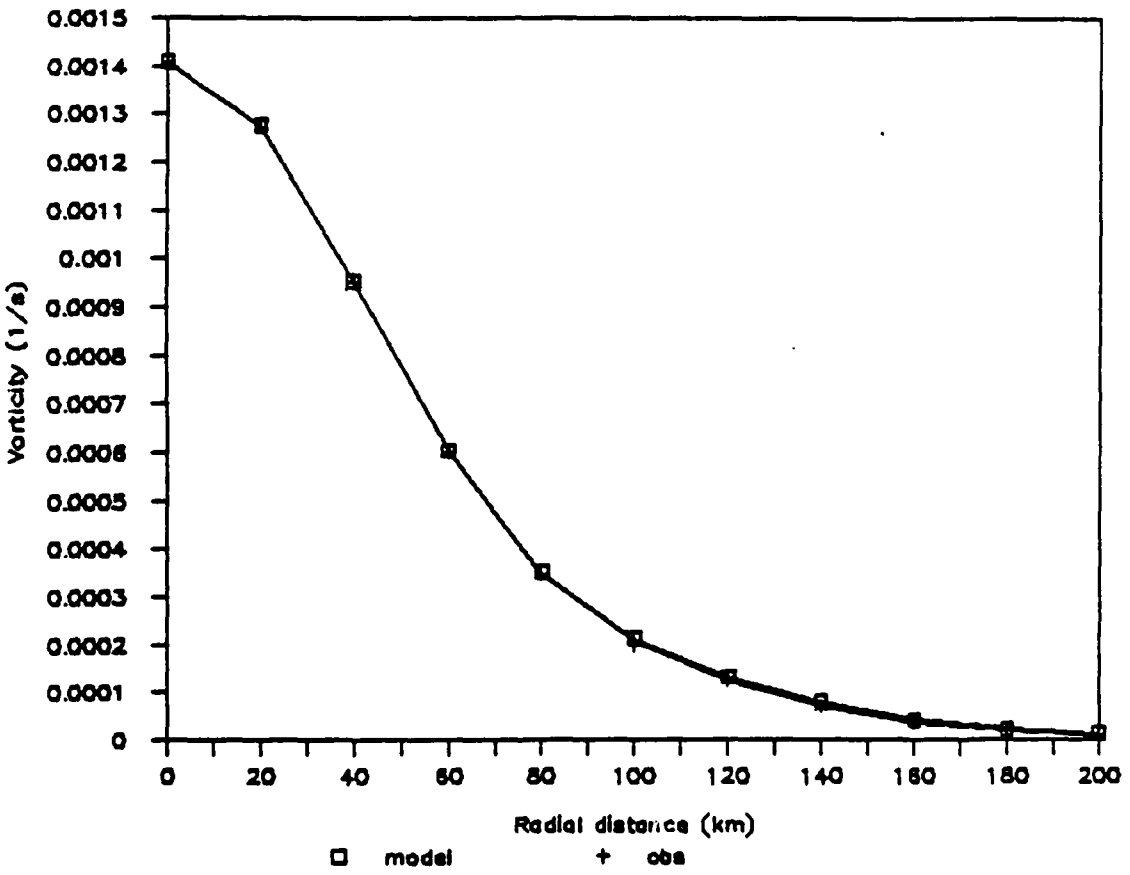


Figure 4.14: The radial profile of the vortices at 48 hours, comparing the observed and model vortices.

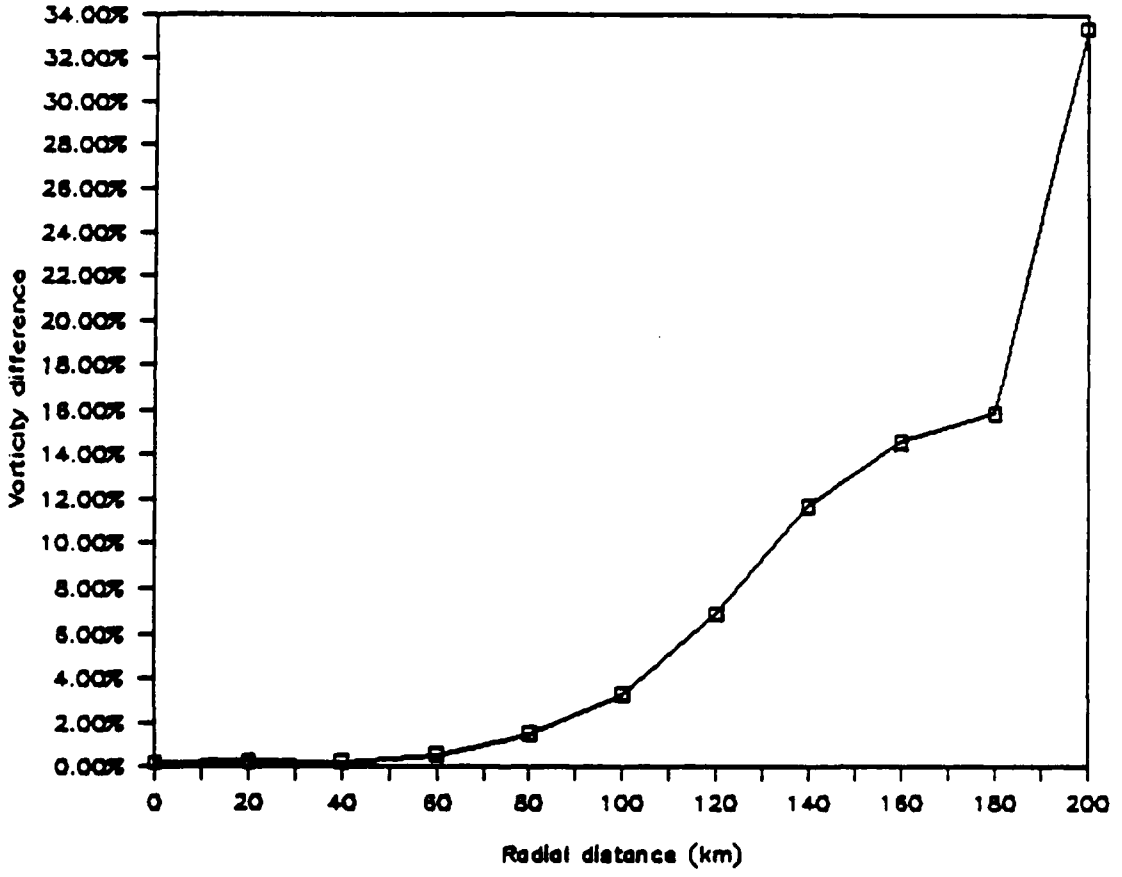


Figure 4.15: The precentage difference in the average radial vorticity between the observed and model vortices at 48 hours.

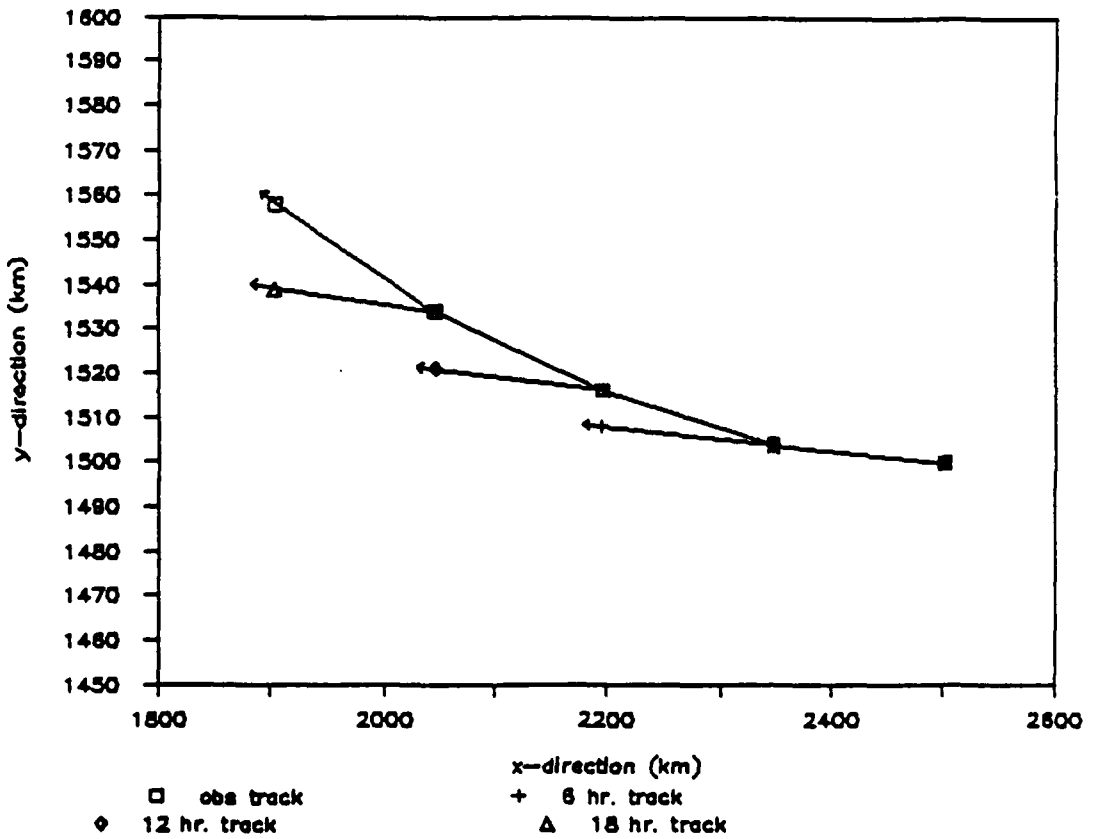


Figure 4.16: This diagram shows the observed track, starting at $x = 2500$ km, $y = 1500$ km, from 0 to 24 hours. Also shown are the first 6 hours of the model tracks started where the observed vortex is located, at 6, 12, and 18 hours.

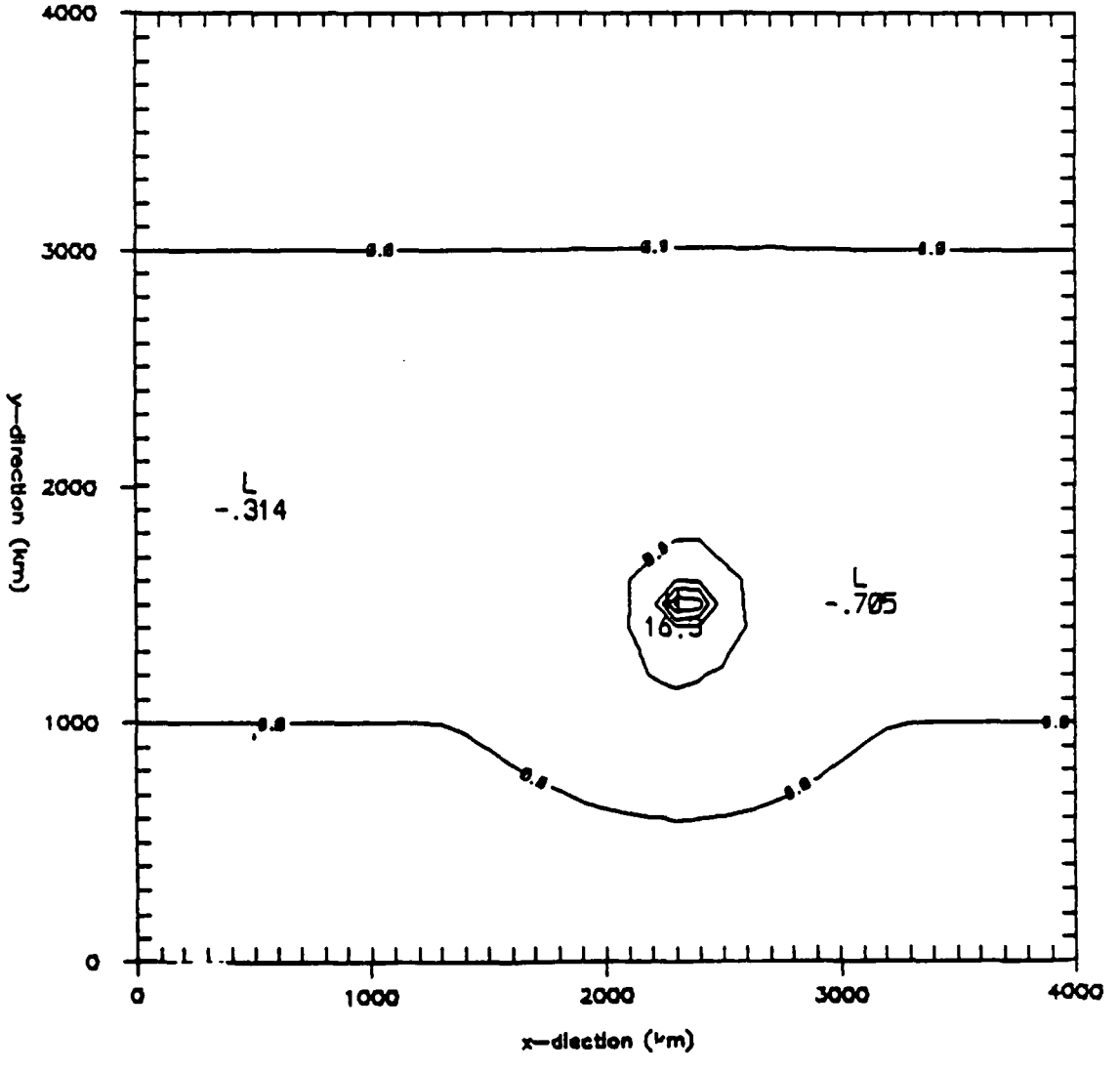


Figure 4.17: The normalized vorticity field at 6 hours for the model run started at x = 2500 km, y = 1500 km. We denote this as the (6 hr) initial observed vorticity field, with the vortex center at x = 2347 km, y = 1504 km.

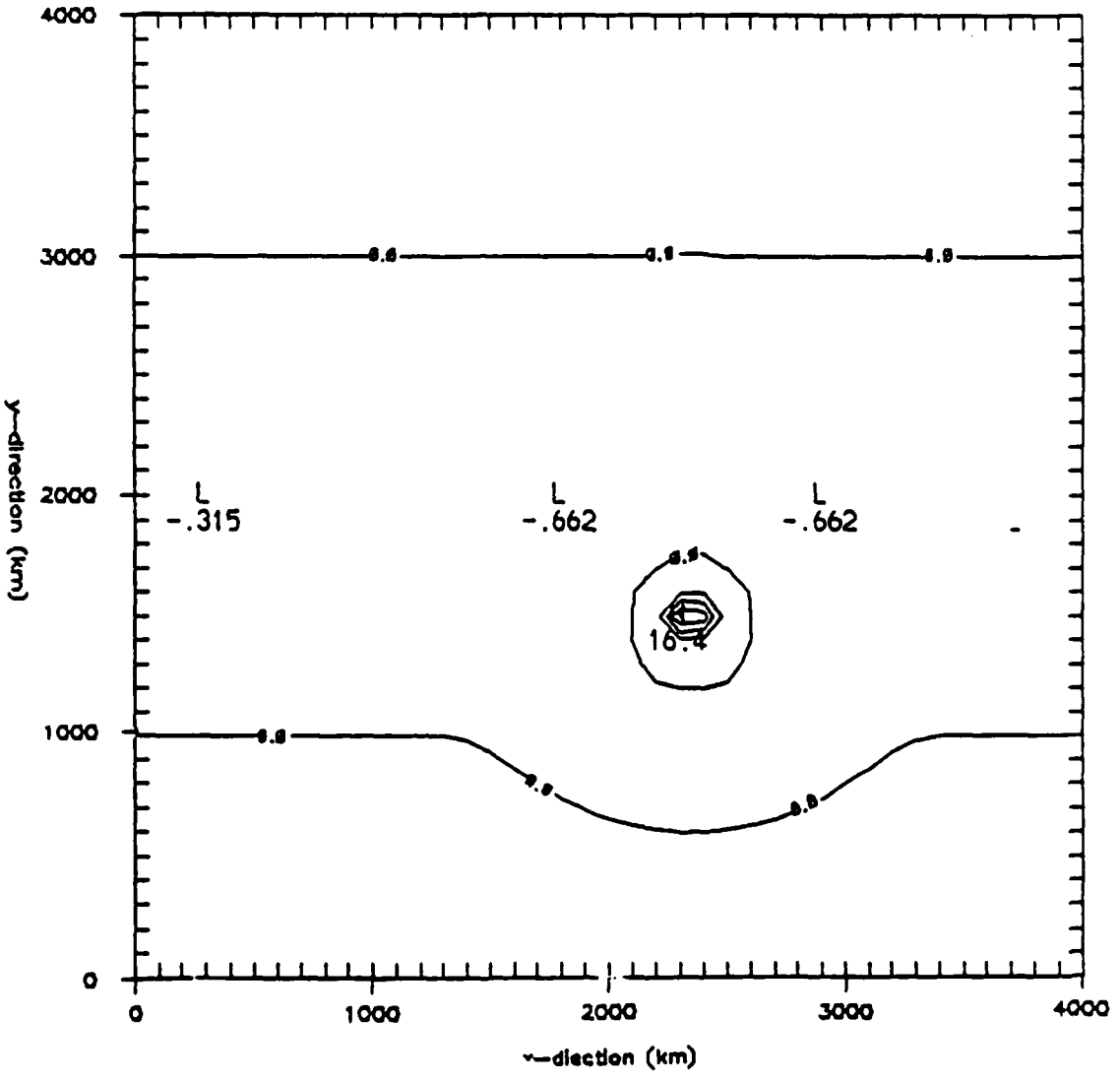


Figure 4.18: The initial normalized vorticity field for the (6 hr) model run, with the vortex centered at $x = 2347$ km, $y = 1504$ km.

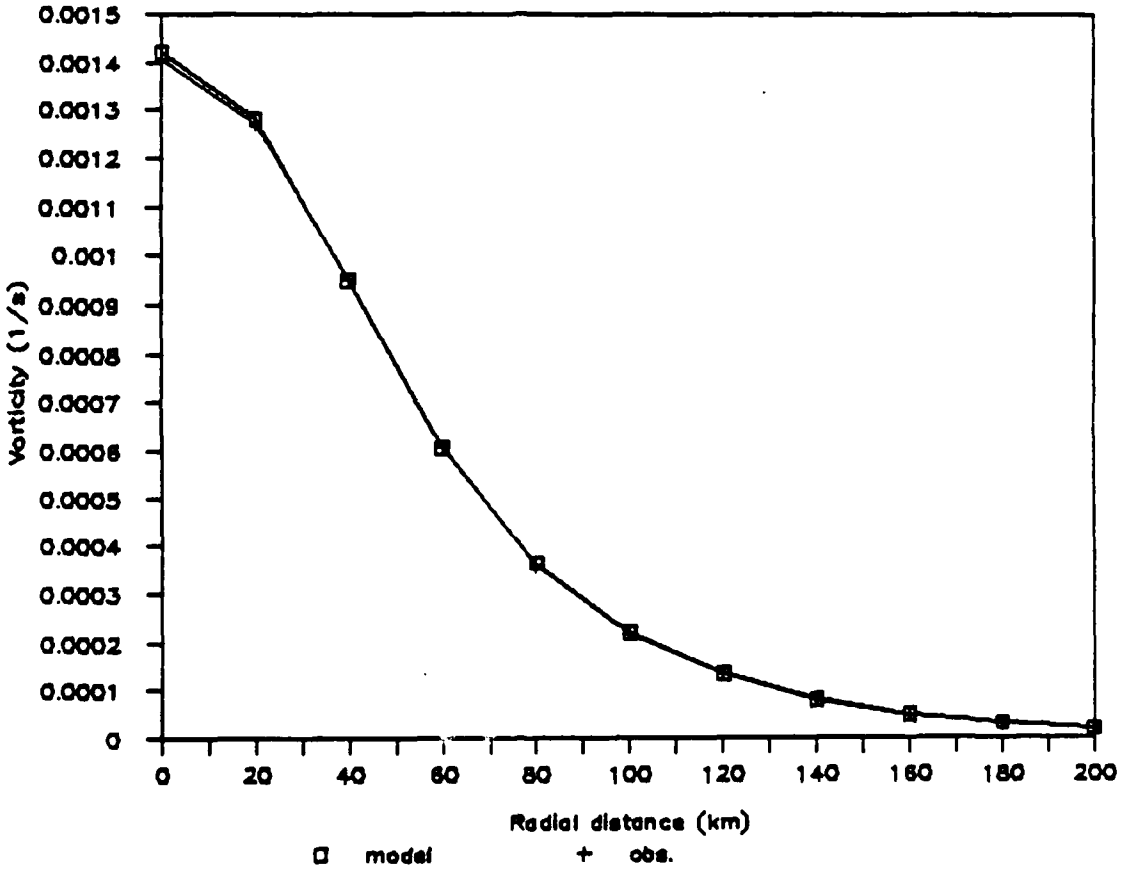


Figure 4.19: The radial profile of the vortices at the (6 hr) initial time. This figure compares average radial vorticity of the observed and model vortices.

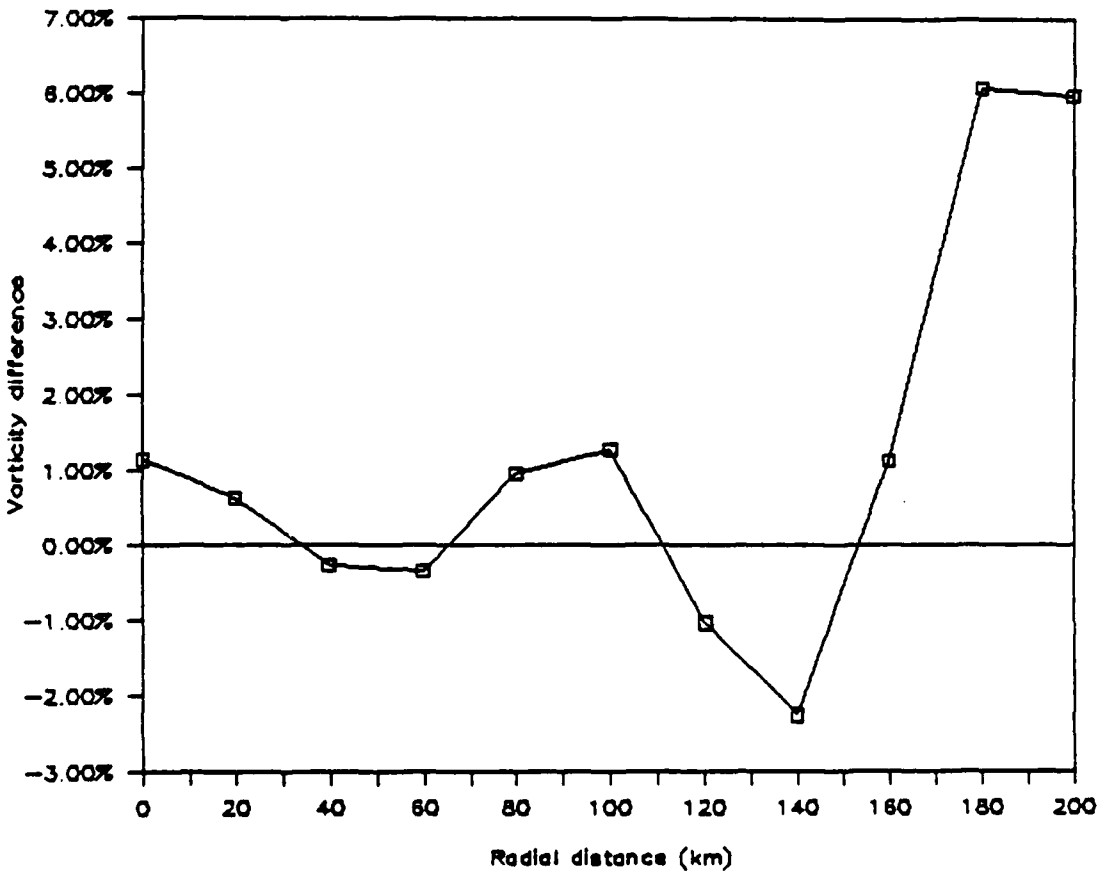


Figure 4.20: The percentage difference in the average radial vorticity between observed and model initial (6 hr) vortices.

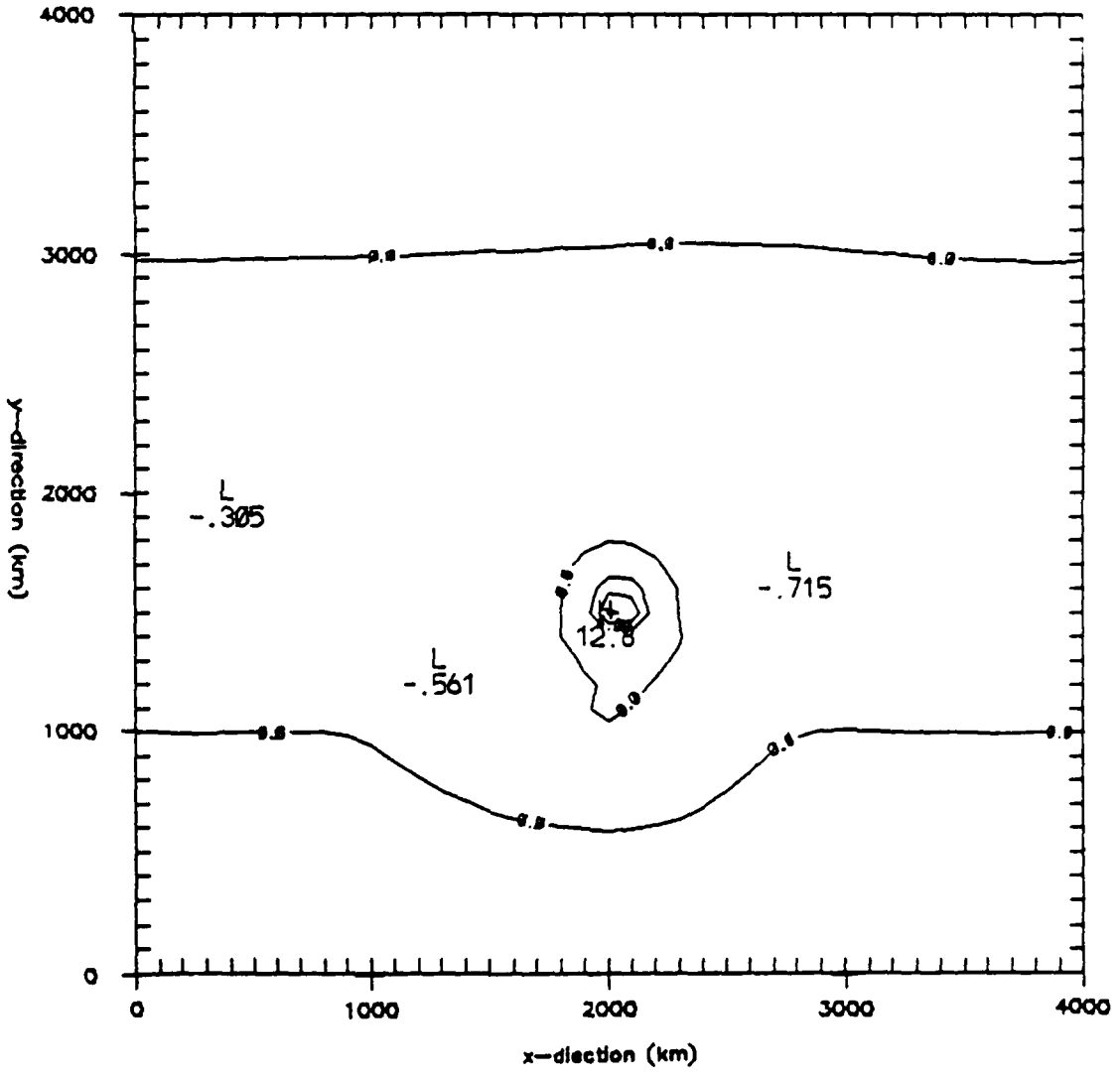


Figure 4.21: The normalized vorticity field at 18 hours for the model run started at x = 2500 km, y = 1500 km. We denote this as the (18 hr) initial observed vorticity field. with the vortex center at x = 2047 km, y = 1534 km.

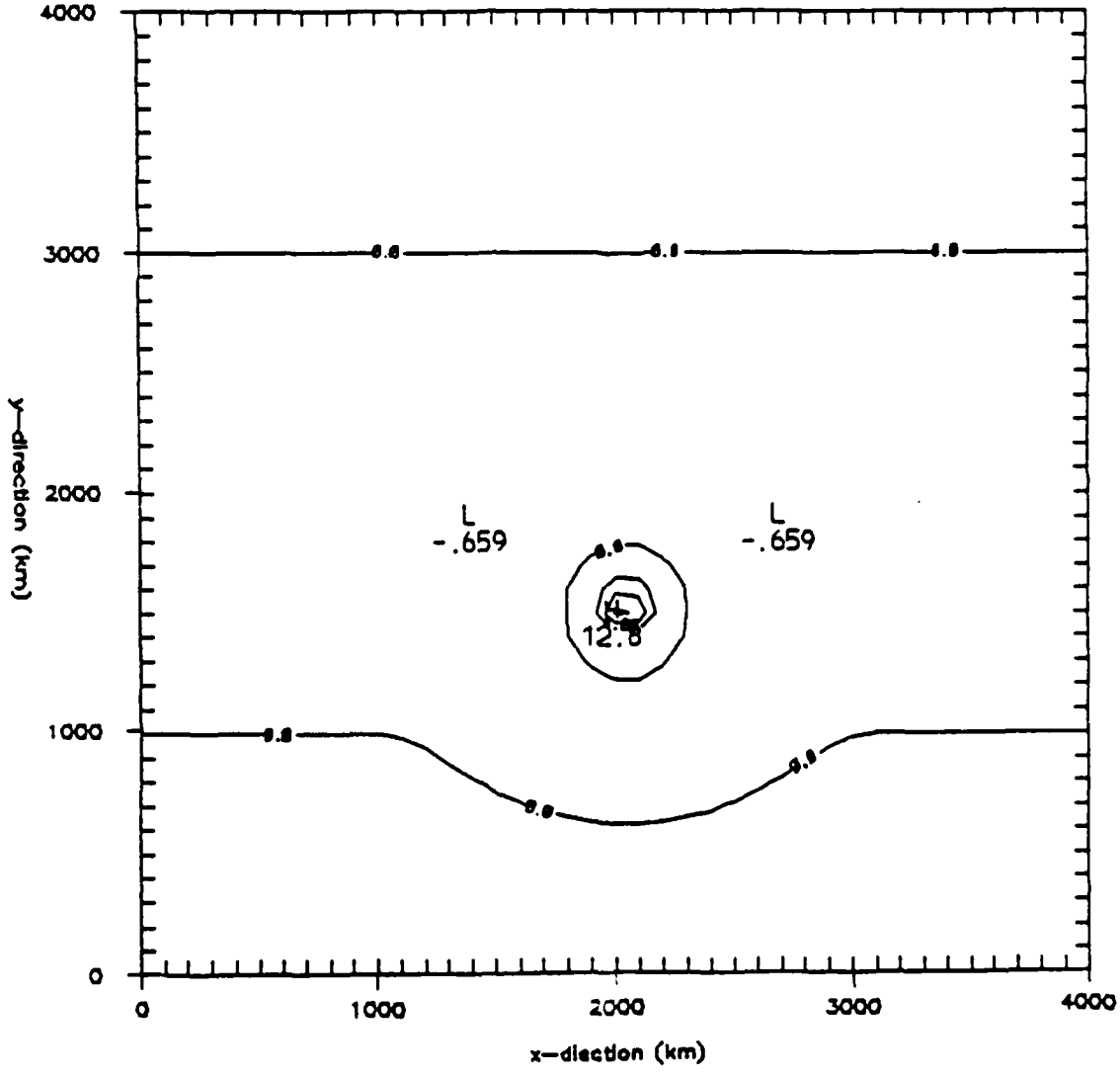


Figure 4.22: The initial normalized vorticity field for the (18 hr) model run at 18 hours, with the vortex centered at x = 2047 km, y = 1534 km.

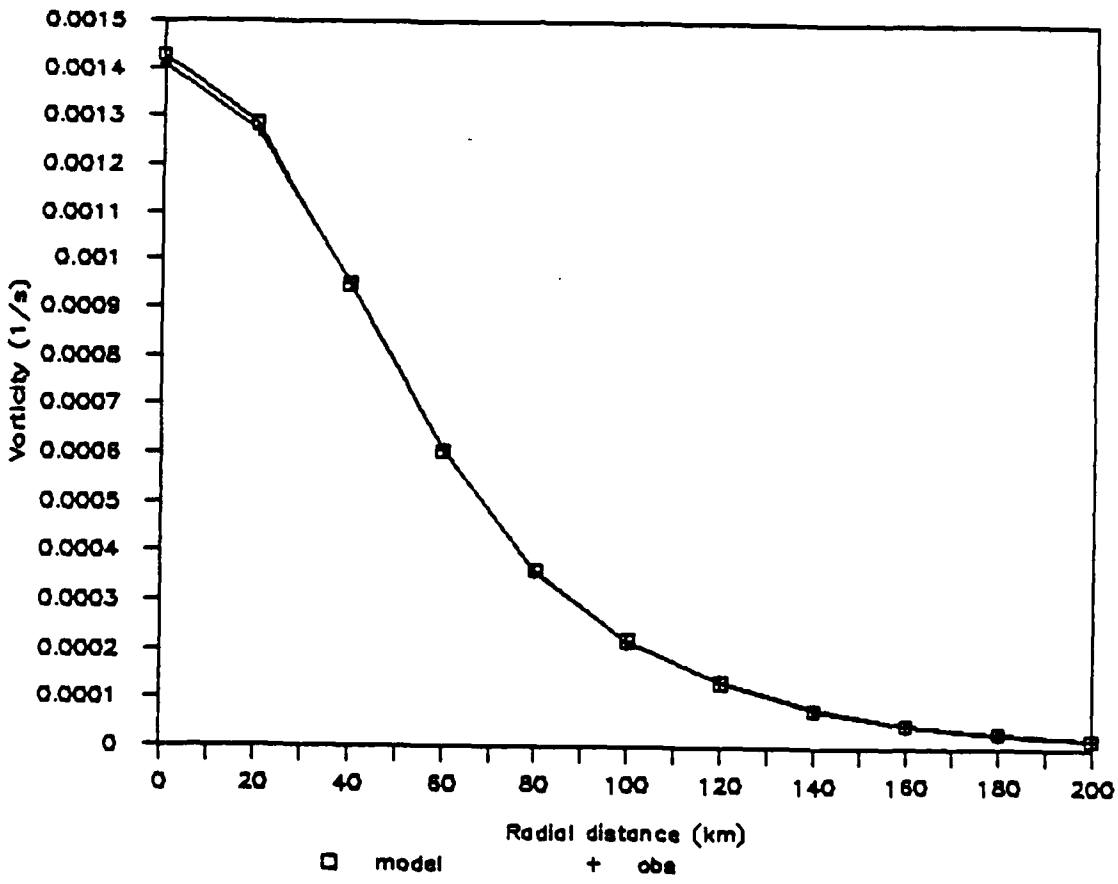


Figure 4.23: The radial profile of the vortices at the (18 hr) initial time. This figure compares average radial vorticity of the observed and model vortices.

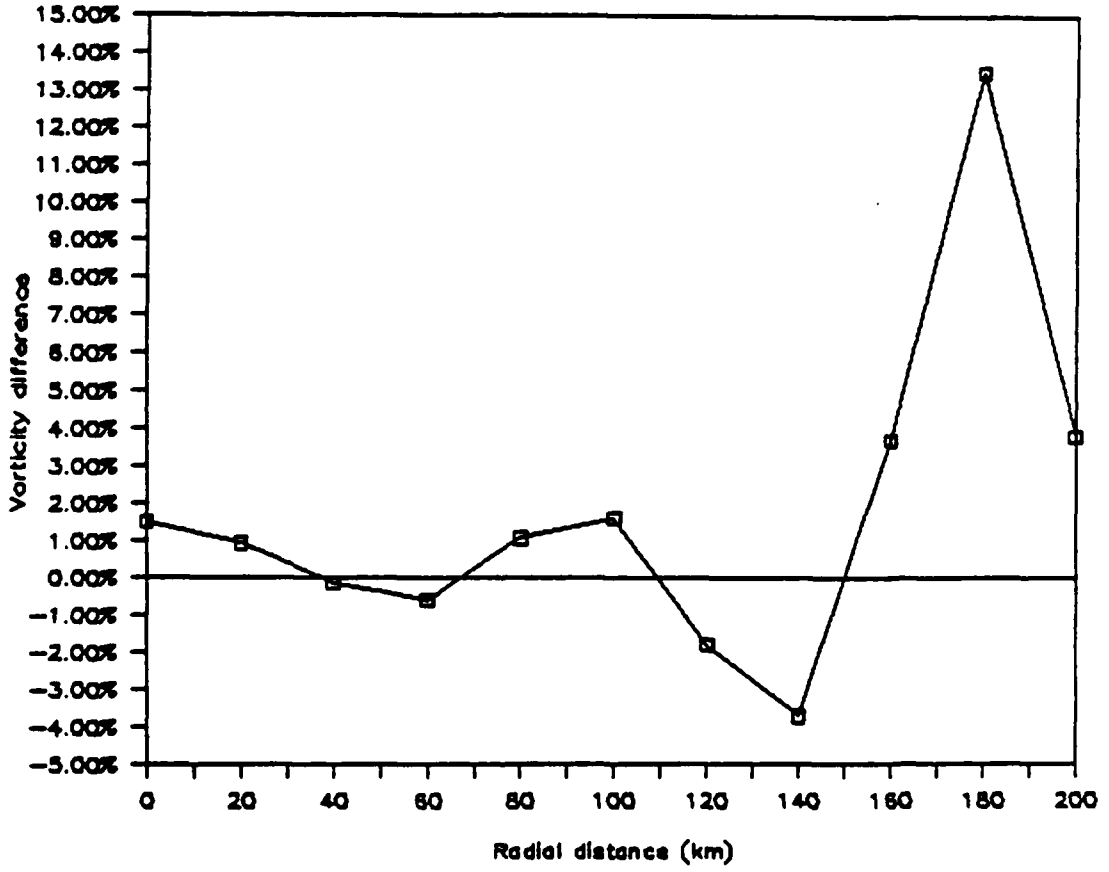


Figure 4.24: The precentage difference in the average radial vorticity between observed and model initial (18 hr) vortices.

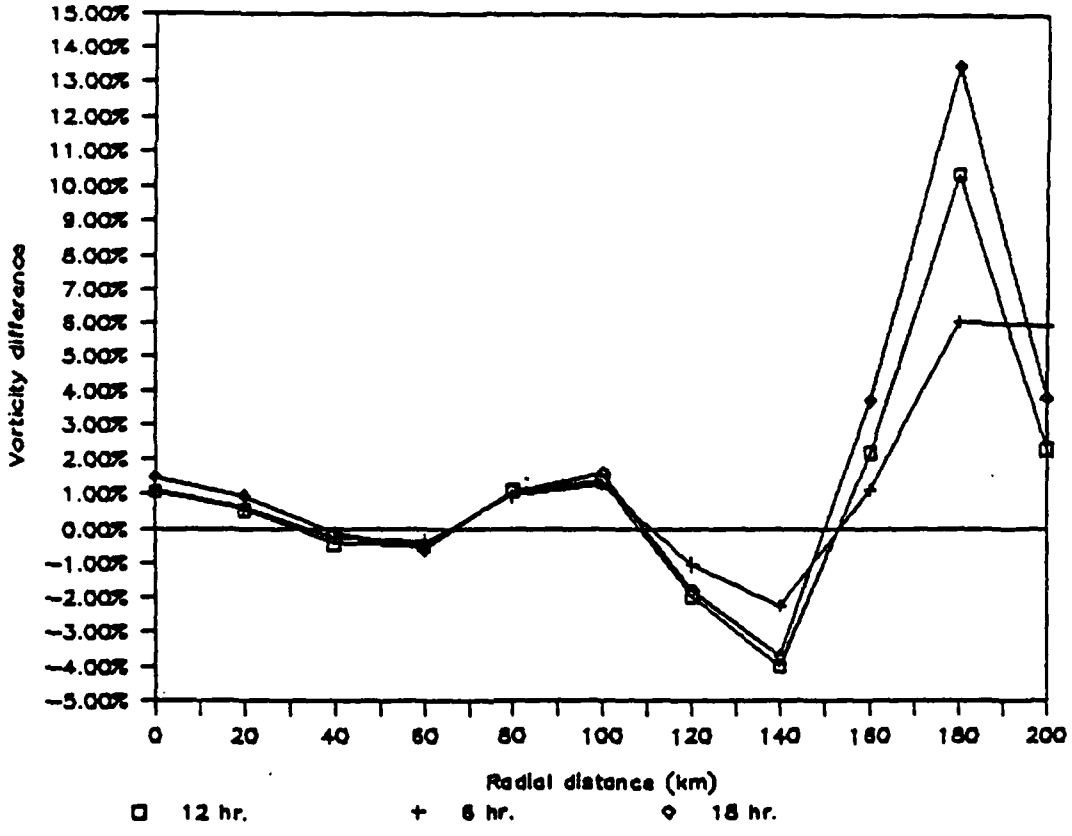


Figure 4.25: A comparison of the percentage difference in the average radial vorticity between observed and model initial vortices, at 6, 12, and 18 hours.

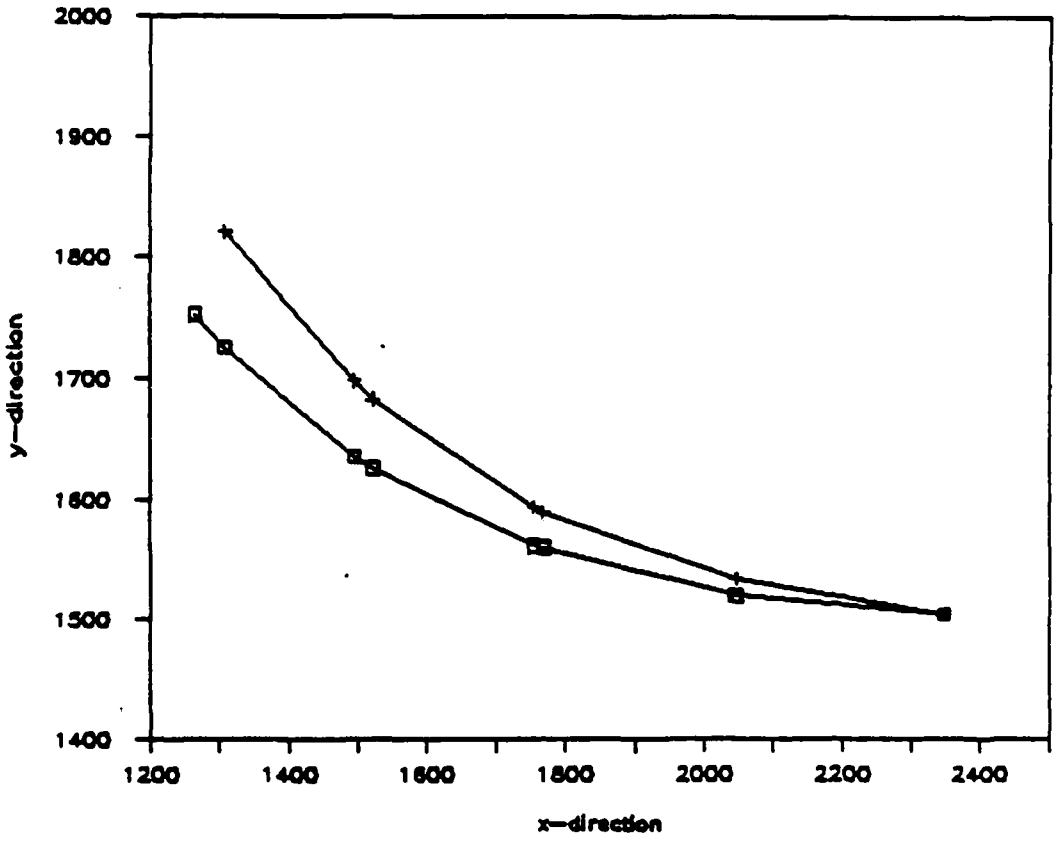


Figure 4.26: The observed and model vortex tracks from 0 to 48 hours, for the (6 hr) initial vortex.

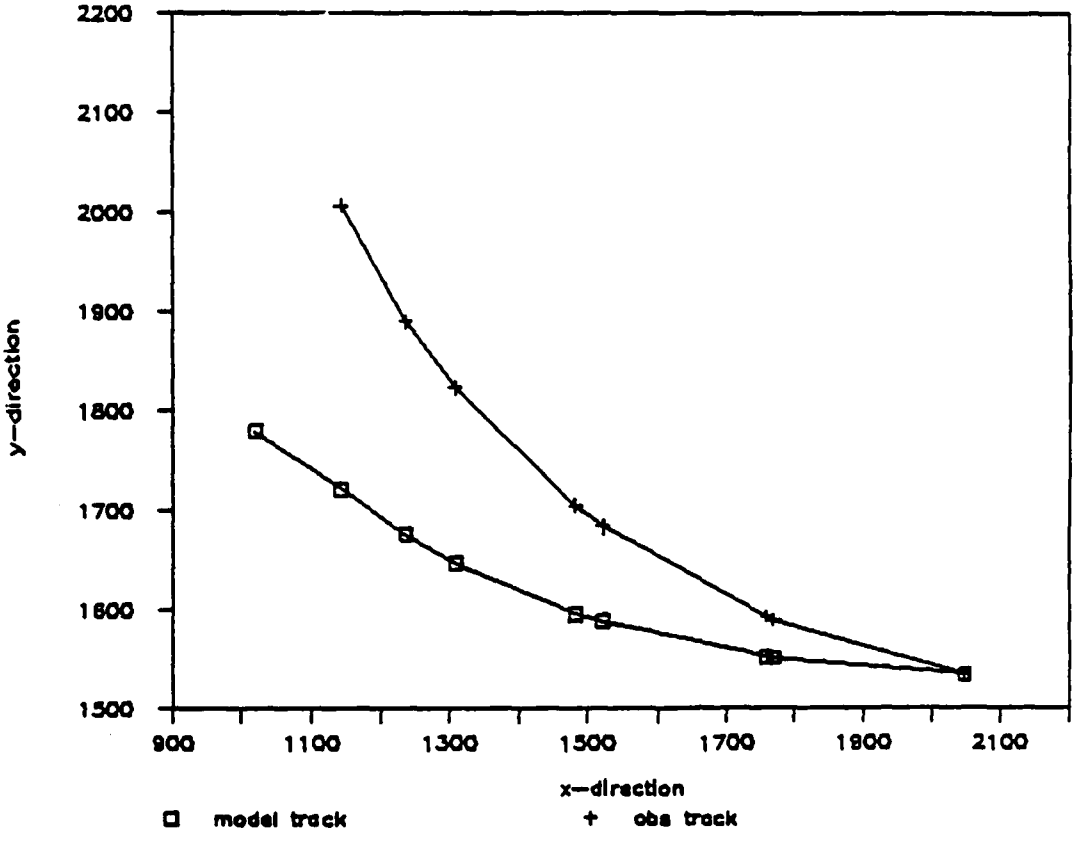


Figure 4 27: The observed and model vortex tracks from 0 to 48 hours, for the (18 hr) initial vortex.

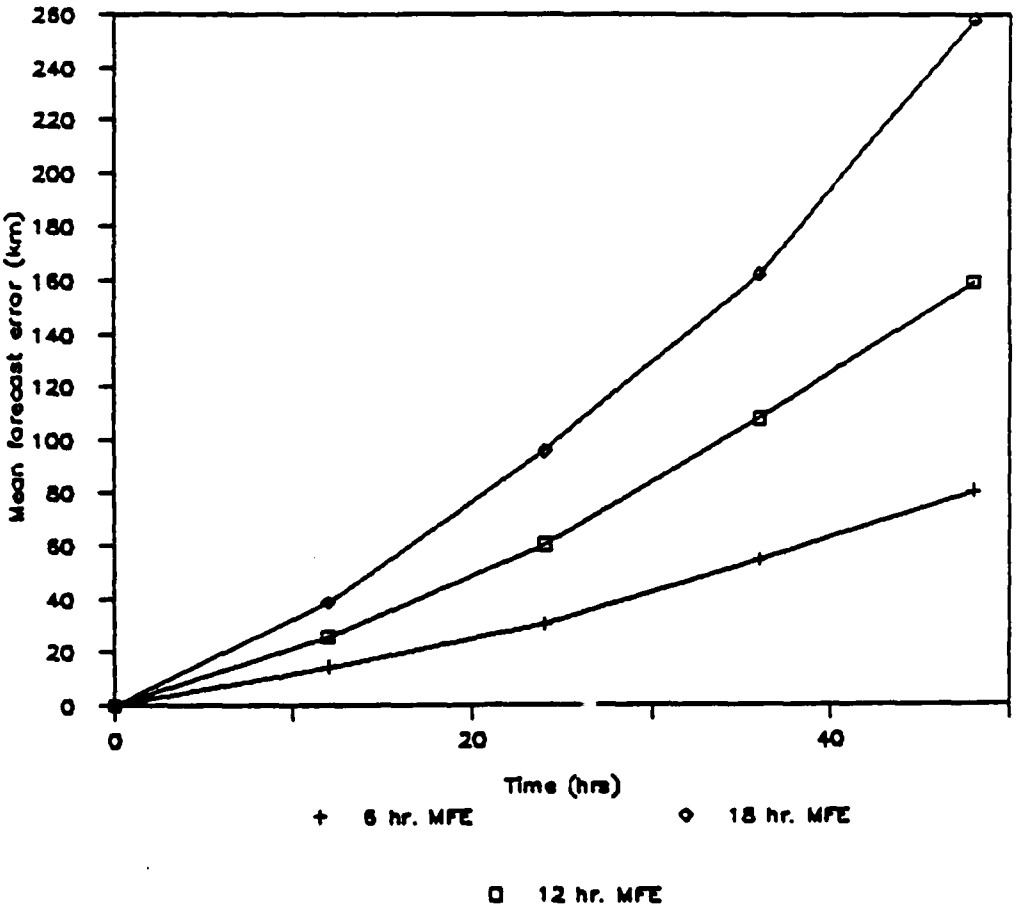


Figure 4.28: A comparison between the 6, 12, and 18 hour mean forecast error of the model vortex track.

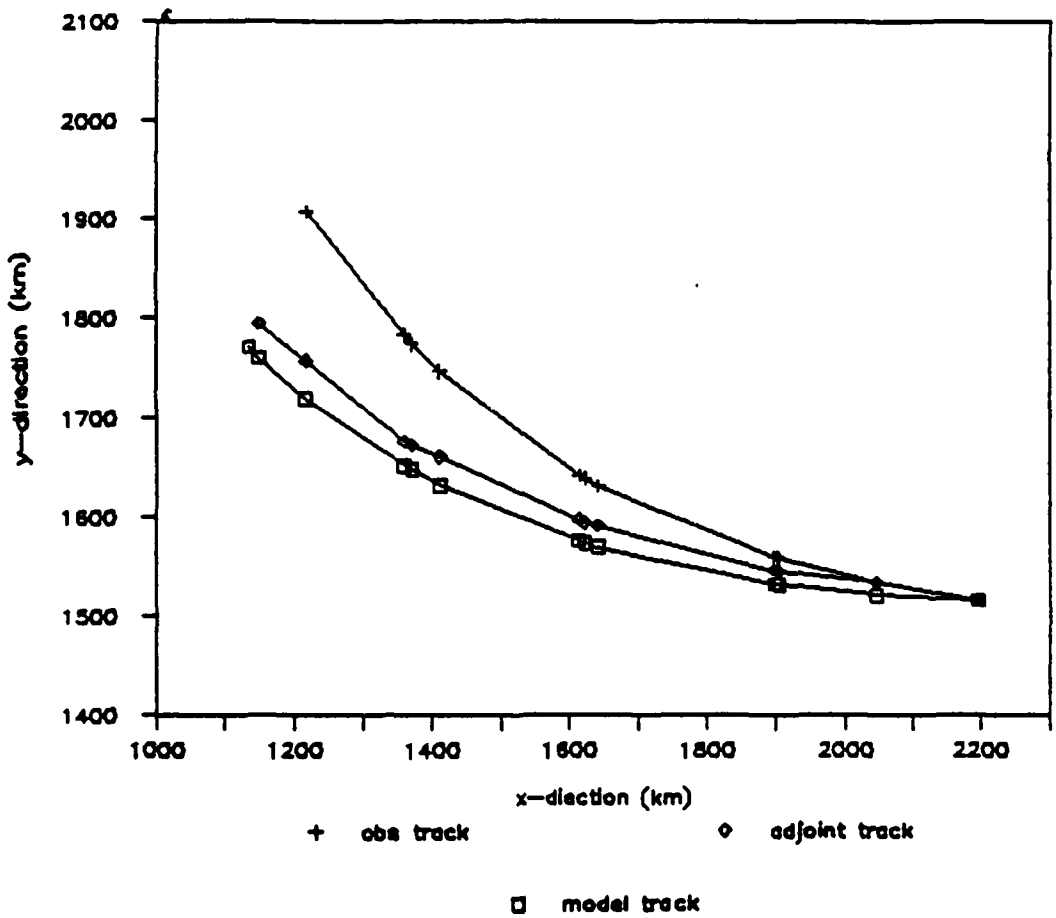


Figure 4.29: The comparison between a simulated hurricane track using the adjoint method and the observed and model tracks. In the case the adjoint track starts to deviate from the observed track 6 hours after the initial time.

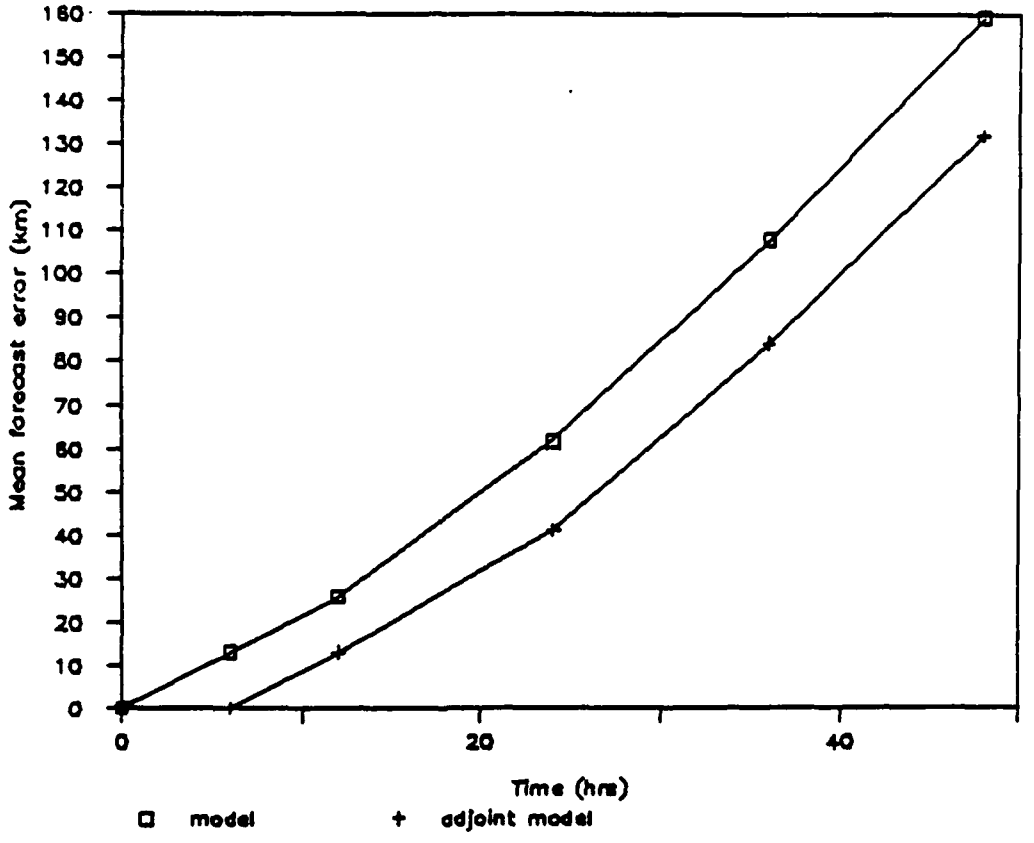


Figure 4.30: The mean forecast error of the model track and the adjoint track for the 6 hour case.

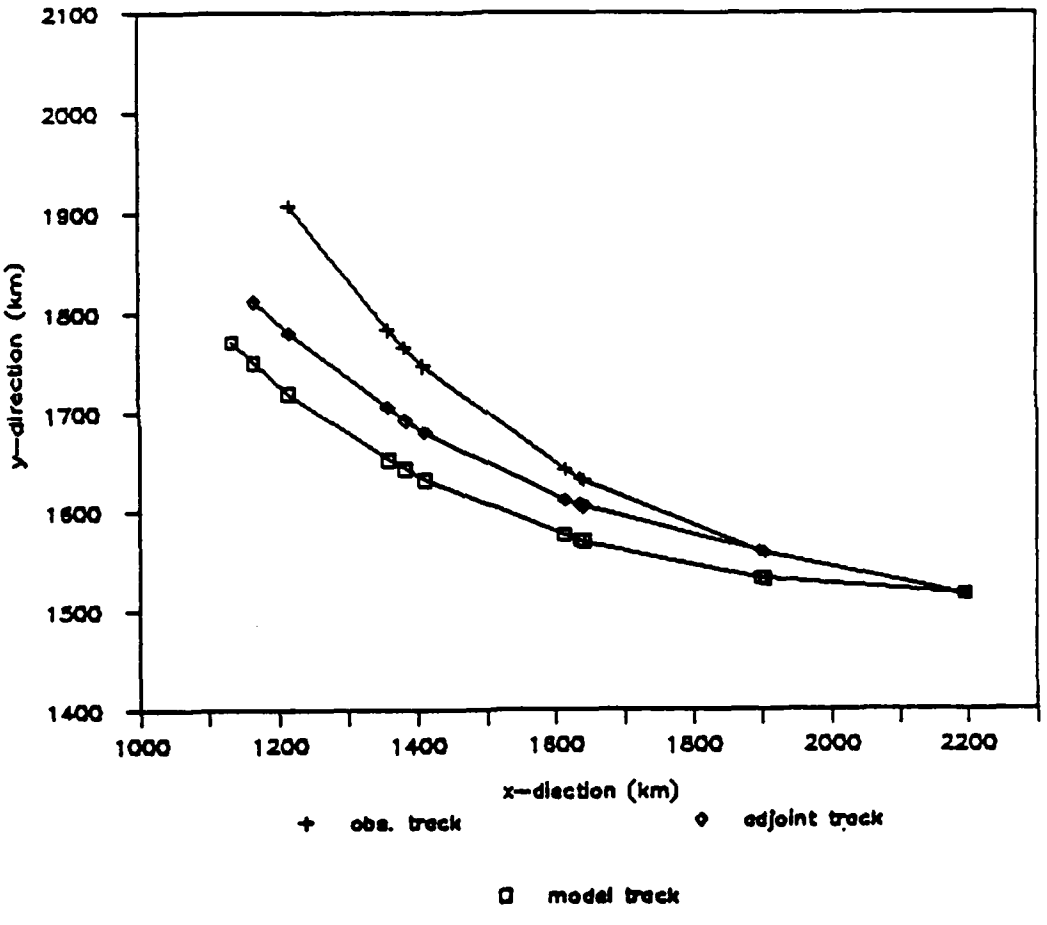


Figure 4.31: The comparison between a simulated hurricane track using the adjoint method and the observed and model tracks. In the case the adjoint track starts to deviate from the observed track 12 hours after the initial time.

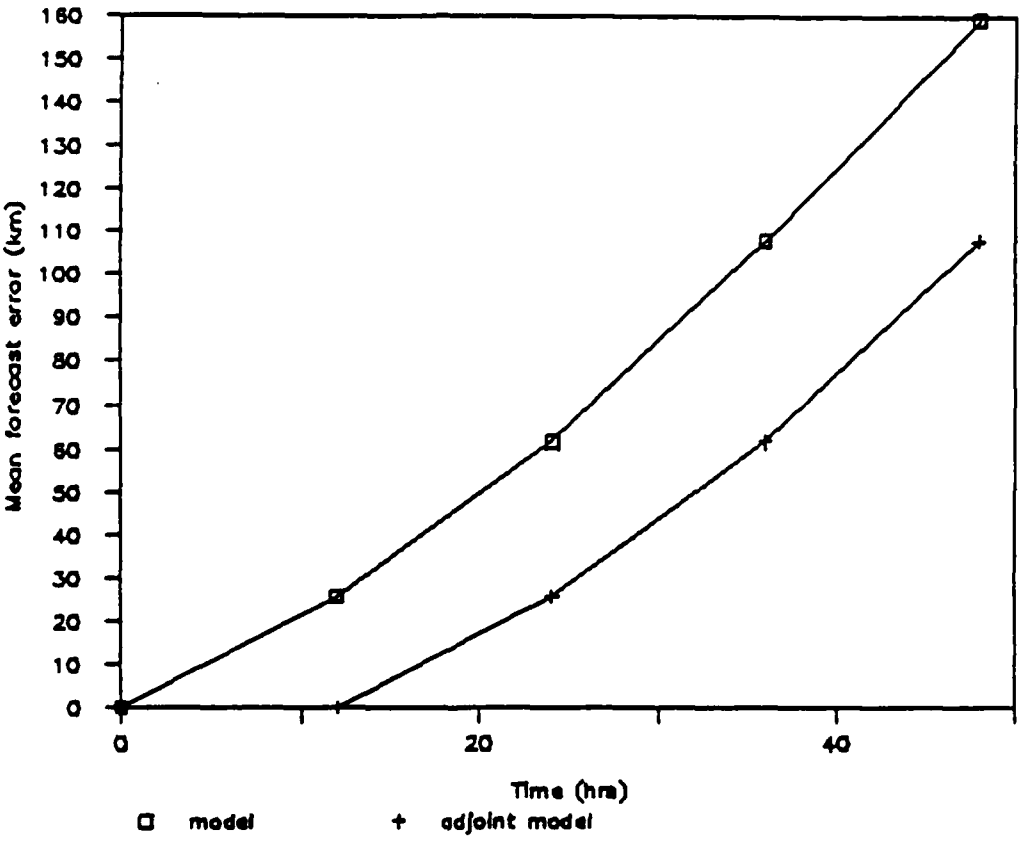


Figure 4.32: The mean forecast error of the model track and the adjoint track for the 12 hour case.

Chapter 5

SUMMARY AND CONCLUSIONS

In this study, we have explored the use of the adjoint method to improve hurricane track forecasts. This is done by successive forward-backward integrations using the adjoint method, which modify the model's initial conditions. The result of each forward-backward integration is the gradient of the distance function, where the distance function is related to the scalar which measures the distance between the observed and model forecast hurricane track. The gradient of the distance function is then used in a minimization scheme which will modify the initial conditions. These new initial conditions will produce a model track closer to the observed track.

The adjoint method developed by Talagrand and Courtier (1987), was derived in this study using the nondivergent barotropic model. Both the nondivergent vorticity equation and the adjoint equation are solved on a β-plane using the spectral method and the Adams-Bashforth time differencing scheme. In chapter 2 the theory of the adjoint method and the nondivergent model are described. The adjoint equation of the nondivergent barotropic model is also derived in chapter 2. In the study by Talagrand and Courtier (1987), the adjoint equation is derived from the nondivergent vorticity equation using spherical coordinates. But here we derive the adjoint equation using the nondivergent vorticity equation on a β-plane. A stability argument developed by Talagrand (1981a), discussed in chapter 2, showed that the adjoint method is the proper approach for a forward-backward assimilation. Also discussed in this chapter are minimization schemes which can be used to reduce the distance function, and in turn modify the initial conditions.

In chapter 3 the nondivergent vorticity equation and the adjoint equation are discretized in space and time, using the spectral method and the Adams-Bashforth scheme.

respectively. The spectral method is used because of its simplicity and accuracy. The Fourier-Galerkin method used here does not produce aliasing error, and the linear terms of the equations have a very simple form in spectral space. However, as discussed in DeMaria (1983), there are disadvantages in using this method. First is the need to assume periodicity in both the east-west and north-south directions. Periodicity in the north-south direction is not appropriate to use on an equatorial β-plane since the governing equation contains operators which are not periodic for this case. The periodic boundary conditions also make it difficult to use real data for initial conditions. So, it is doubtful if this method could be used in an operational model. The time differencing scheme used in this study is different from the one used by Talagrand and Courtier. In their derivation the leapfrog method was used, while we used the Adams-Bashforth method in our derivation. We use the Adams-Bashforth method because this method eliminates the computational mode that is sometimes troublesome in the leapfrog method.

In chapter 4 we ran experiments using the nondivergent model to indicate how the adjoint method can improve hurricane track forecasts. First, the model was integrated forward using slightly different initial vortices started from the same position. The results from this experiment show that small changes in the vortex structure, produce large changes in hurricane track. This indicates that there is important information in the vortex track itself, and that the adjoint equation could be very useful in improving the track forecast. In the second experiment, we test how sensitive the hurricane track is to changes in the vortex structure. We compared the model track in the first experiment with two model tracks started at the positions where the observed track was at 6 hours and 18 hours. These slightly different vortices produce expected results. A greater difference in the vortex structure produced a larger mean forecast error in the hurricane track. These experiments show that the vortex structure strongly influences the hurricane track, and also indicates that the adjoint method can be very powerful in improving the hurricane track forecast.

Although the adjoint method shows promise, there are problems with it. First, the derivation of the adjoint equation is complicated, and would become much more so for

more complicated, realistic models. Also for several reasons the adjoint method consumes much more computer time than the original model. One reason is because the forward-backward integrations are required in the adjoint method. Again, as the complexity of the model increases, the amount of computer time required for each forward-backward integration should significantly increase. The amount of computer time also increases using the descent algorithm. The descent algorithm may also, depending on the one used, take up large amounts of memory also. However, in the future, with the continuing advances in the speed and amount of memory of computers, these problems can be reduced.

This study was to show how useful the adjoint method can be in forecasting hurricane tracks. From here there are many possibilities for further study - first, running this model with the adjoint method, and studying the improvement in the forecast hurricane tracks. From here, more complicated and realistic models using the adjoint method can be developed. This step might involve the use of Chebyshev spectral methods, instead of the double Fourier series used here. For limited-area models, Chebyshev polynomials are capable of handling general boundary conditions that the double Fourier method can not (Fulton and Schubert, 1987). Another possibilities could include using spherical coordinates, and vertically layered models, such as the one described by DeMaria (1983). Also different time differencing schemes, such as the fourth-order Runge-Kutta method, might be studied. Also studies should be done to determine the best way to reduce the noise developed using the adjoint method, which was discussed in chapter 2. Is a penalty term the best way to do this? And if so, what is the best one to use? Another question is how to integrate actual observations into the adjoint method. The use of this method is new and very promising, and should be studied in the future.

REFERENCES

Birchfield, G.E., 1960: Numerical prediction of hurricane movement with the use of a fine grid. *J. Meteor.*, **17**, 406–414.

Buckley A., and A. Le Nir, 1983: QN–like variable storage conjugate gradients. *Mathematical Programing*, **27**, 155–175

Conte S., and C. de Boor, 1981: *Elementry Numerical Analysis: An Algorithmic Approach*. McGraw–Hill International Book Company, Auckland and others.

Courtier, P., and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Quart. J. Roy. Meteor. Soc.*, **113**, 1329– 1347.

DeMaria, M., 1983: *Experiments with a Spectral Tropical Cyclone Model. Atm.* Sci. Paper No. 371, Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado.

DeMaria, M., 1984: Experiments with a Spectral Tropical Cyclone Model. *J. Atm. Sci.*, **41**, No 5, 901–924.

DeMaria, M., 1985: Tropical Cyclone Motion in a Nondivergent Barotropic Model. *Mon. Wea. Rev.*, **113**, 1199–1210.

DeMaria, M., 1987: Tropical Cyclone Track Prediction with a Barotropic Spectral Model. *Mon. Wea. Rev.*, **115**, 2346–2357.

Fulton S.R., and W.H. Schubert, 1987: Chebyshev Spectral Methods for Limited- Area Models. Part I.: Model Problem Analysis. *Mon. Wea. Rev.*, **115**, 1940–1953.

Ghil, M., S.E. Cohn, and A. Dalcher, 1982: Sequential estimation, data assimilation and initialization. *The interaction between objective analysis and initialization* (Editor D. Williamson). Proceedings of the Fourteenth Stanstead Seminar, McGill University, Montreal Canada, 83–97.

Gill, P.E., W. Murray, and M.H. Wright, 1982: *Practical optimization.* Academic Press. London.

Gottlieb, D. and S.A. Orszag, 1977: *Numerical Analysis of Spectral Methods: Theory and Applications.* Regional conference series in applied mathematics, No. 26. Haltiner G., and R. Williams, 1980: *Numerical Prediction and Dynamic Meteorology.* John Wiley and Sons, New York, Chichester, Brisbane, and Toronto.

Kasahara, A., 1957: The numerical prediction of hurricane movement with the barotropic model. *J. Meteor.*, **14**, 386–402.

Krylov, V.I., 1962: *Approximate Calculation of Integrals.* The Macmillan Company, New York, 357 pp.

Lions, J.L., 1971: *Optimal control of systems governed by partial differential equations* (English translation). Springer-Verlag, Berlin.

Neumann, C., and J. Pelisser, 1981: Models for the Prediction of Tropical Cyclone Motion over the North Atlantic: An Operational Evaluation. *Mon. Wea. Rev.*, **109**, 522–538.

Phillips, N.A., 1959: An example of nonlinear computational instability. *The Atmosphere and the Sea in Motion*, Rossby Memorial Volume, New York, Rockefeller Instit. Press, 501–504.

Reed, M., and B. Simon, 1980: *Methods of modern mathematical physics. I: Functional analysis.* Academic Press, New York, U.S.A.

Ralston, A., 1965: *A First Course in Numerical Analysis.* McGraw-Hill Book Company. New York, St. Louis, San Francisco, Toronto, London, and Sydney.

Sanders, F., and R.W. Burpee, 1968: Experiments in barotropic hurricane track forecasting. *J. Appl. Meteor.*, **7**, 313–323.

Sasaki, Y., 1955: Barotropic forecasting for the displacement of typhoon. *J. Meteor. Soc. Japan*, **33**, 1–8.

Simpson R., and H. Riehl, 1981: *The Hurricane and Its Impact.* Louisiana State University Press, Baton Rouge and London.

Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311–1328.

Temperton, C., 1983a: Self-sorting mixed-radix fast Fourier transforms. *J. Comp. Phys.*, **52**, 1–23.

Temperton, C., 1983b: A note on prime factor FFT algorithms. *J. Comp. Phys.*, **52**, 198–204.

Temperton, C., 1983c: Fast mixed-radix real Fourier transforms. *J. Comp. Phys.*, **52**, 340–350.

Young, J., 1968: Comparative properties of some time differencing schemes for linear and nonlinear oscillations. *Mon. Wea. Rev.*, **96**, 6, 357–364.